

The feasibility of the dual-task paradigm as a framework for a clinical test of listening effort in cochlear implant users

Helen Willis

A thesis submitted in fulfilment of the requirements
for the degree of Doctor of Philosophy
of UCL

Department of Speech, Hearing and Phonetic Sciences,
Division of Psychology and Language Sciences,
University College London

2018

Declaration

I, Helen Willis, confirm that the work presented in this thesis is my own.

Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Abstract

The overall aim of this thesis is to evaluate the feasibility of using the behavioural framework of the dual-task paradigm as the basis of a clinical test of listening effort (LE) in cochlear implant (CI) users. It is hypothesised that, if a primary listening task is performed together with a secondary visual task, performance in the visual task will deteriorate as the listening task becomes harder. This deterioration in secondary visual task performance can then provide an index of LE.

An initial series of six experiments progressively modified the dual-task design (in an attempt to optimise its sensitivity to LE), leading to the selection of British English Lexicon (BEL) sentences for the listening task and a digit stream visual task. A further three experiments applied this dual-task to 30 normal hearing (NH) participants listening to normal speech, 30 NH participants listening to CI simulations, and 25 CI users listening through their speech processors. Performance in quiet conditions was compared to that in different levels of background noise. Adaptive tracking procedures were used in an attempt to ensure that the challenge of noise was equal for all participants. This principle was also applied to equalise difficulty in terms of the number of channels used in the spectral resolution of the CI simulations.

As expected, NH participants only exhibited significant deterioration in visual accuracy when noise was present ($p < .001$), suggesting increased LE. Interestingly, however, when CI simulations were applied, this significant visual deterioration occurred immediately in quiet ($p < .001$). The same result occurred in quiet for the CI users too ($p < .001$). Therefore, it appears that the degraded auditory input provided by CI induces LE even in optimal listening conditions.

These results suggest that the dual-task paradigm could feasibly become a framework for developing a clinical test of LE in the CI user population.

Acknowledgements

The completion of this doctorate has only been possible because of the culmination of the kindness, generosity and expertise of a whole host of individuals.

Firstly, I would like to thank my primary supervisor, Professor Stuart Rosen. Stuart was able to transform my enthusiastic, but mediocre, research proposal into one of sound scientific integrity. Without Stuart, the genesis of this thesis would not have been possible. His ongoing supervision throughout the four years of this PhD, never failing to provide astute and constructive criticism and feedback, then ensured that I avoided many of the abundant pitfalls inherent in research. Not only this, Stuart showed magnanimity with his time by always being willing to meet with me in person whenever I needed to confer with him (in spite of his undoubtedly hectic schedule of lecturing and research responsibilities). As a result, I am absolutely beholden to Stuart for his support.

I would also like to show appreciation to both Dr. Deborah Vickers and Dr. Tim Green for readily agreeing to be my secondary supervisors. Debi was indispensable particularly in the early stages of my PhD, providing sage advice about the challenges and obstacles involved in cochlear implant research. Tim has been essential in enabling this doctoral research to gain momentum after my pilot study; I am particularly indebted to Tim for his munificence in both time and effort in providing me with bespoke MATLAB programming, without which none of my subsequent experiments could possibly have proceeded. In addition to this, he was always fantastically quick in responding to any queries, or requests, for help with the execution of these programming scripts, saving many participants and myself the trauma of stress and delay within the testing sessions. I am sincerely grateful for his consideration at these pivotal moments.

I would also like to thank Professor Alessandro Leidi, the founding director of Statistical Support Services, for his patient instruction and astute advice related to the design of the inferential statistical analyses, construction of all the necessary SPSS syntax and execution of power analysis (using PASS). The support provided by Sam Hutton, Research Support Specialist at SR Research, too was critical; his provision of tailored beta software and superb one-to-one training with the eyetracker equipment enabled smooth operation of pupillometry for my pilot study.

Another individual I would like to thank is Andrew Clark, within the Department of Speech, Hearing and Phonetics at UCL. His expertise with technology has been absolutely invaluable and he was ceaselessly tolerant of my repeated approaches to him whenever I required his assistance in the earliest stages of my piloting, or for technical issues such as calibration and equipment loan.

Regarding the Department of Speech, Hearing and Phonetic Sciences, I would also like to thank the PhD students within this community, as well as those from across the Division of Psychology and Language Sciences at UCL. Everyone has been unfailingly warm and welcoming from the first day, and they have become a critical forum for advice for all academia-related matters, as well as a source of friendship.

Another cohort of individuals I must thank are all my participants, both deaf and hearing. Their data form the fundamental backbone of this work and I am touched by, and profoundly grateful for, their extraordinary level of conscientiousness and eagerness to devote collectively hundreds of hours to this doctoral research.

I would also like to pay tribute to my mother and father. They provide an absolutely vital touchstone and anchor, and I cannot express enough gratitude for their belief and support.

Finally, I cannot forget to thank all my benefactors at *Action On Hearing Loss* and *Cochlear UK* who provided the generous studentship to fund this doctoral research. I feel honoured and lucky to have been given this opportunity. It is my hope that this data will be of value in helping others, like myself, with cochlear implants to flourish and realise their full potential in spite of their hearing impairment.

Contents

PART ONE: Introduction

1 Chapter One: Literature Review

1.1.	Overview of the context.....	22
1.2.	Understanding listening effort: A literature review.....	24
1.2.1.	What is listening effort?.....	24
1.2.2.	Why is listening effort important?.....	32
1.2.2.1.	The effects of stress.....	34
1.2.2.2.	The effects of the connectome disease.....	44
1.2.2.3.	The reality of electric hearing.....	48
1.2.3.	How can listening effort be measured?.....	57
1.2.3.1.	Subjective measures.....	57
1.2.3.2.	Physiological indices.....	60
1.2.3.3.	Behavioural assessments.....	68
1.2.3.4.	Feasibility of LE measurement in clinical environments	74
1.3.	Statement of aim of thesis.....	76
1.4.	Outline for remainder of thesis.....	76

PART TWO: Pilot Study – Recall Dual-Task

2 Chapter Two: Pilot study testing a recall dual-task

2.1.	Introduction.....	79
2.2.	Fundamental principles of the dual-task paradigm	79
2.3.	Hypotheses.....	83
2.4.	Ethics and screening.....	83
2.5.	Pilot Study.....	84
2.5.1.	Methods.....	84
2.5.1.1.	Participants.....	84
2.5.1.2.	Procedure.....	85
2.5.2.	Results and Discussion: Descriptive statistics.....	96

2.5.2.1.	LE measurement (secondary task performance): Visual accuracy.....	96
2.5.2.2.	Evaluation of validity of LE measurement: Pupil response	98
2.5.2.3.	Evaluation of validity of LE measurement: Subjective ratings.....	100
2.5.2.4.	Stability of primary task performance: Auditory accuracy.....	103
2.5.2.5.	Evaluation of viability of LE measurement.....	105
2.6.	Conclusions of Pilot Study.....	106

PART THREE: Digit Stream Dual-Task (Normal Hearing Participants)

3 Chapter Three: Refining and testing a digit stream dual-task (Experiments 1, 2 & 3)

3.1.	Introduction.....	109
3.2.	Proposed solution.....	109
3.3.	Data collection.....	110
3.4.	Experiment 1.....	111
3.4.1.	Hypotheses.....	111
3.4.2.	Methods.....	111
3.4.2.1.	Participants.....	111
3.4.2.2.	Procedure.....	111
3.4.3.	Results and Discussion: Descriptive statistics.....	118
3.4.3.1.	Secondary task performance: Visual accuracy.....	118
3.4.3.2.	Stability of primary task performance: Auditory accuracy.....	120
3.4.3.3.	Evaluation of potential viability of new LE measurement.....	121
3.5.	Conclusions of Experiment 1.....	123
3.6.	Proposed solution.....	123
3.7.	Experiment 2.....	124
3.7.1.	Hypotheses.....	124
3.7.2.	Methods.....	124

3.7.2.1.	Participants.....	124
3.7.2.2.	Procedure.....	124
3.7.3.	Results and Discussion: Descriptive statistics.....	127
3.7.3.1.	Secondary task performance: Visual accuracy.....	127
3.8.	Conclusions of Experiment 2.....	128
3.9.	Experiment 3.....	129
3.9.1.	Introduction.....	129
3.9.2.	Hypotheses.....	129
3.9.3.	Methods.....	129
3.9.3.1.	Participants.....	129
3.9.3.2.	Procedure.....	129
3.9.4.	Results and Discussion: Descriptive and Inferential statistics.....	132
3.9.4.1.	LE measurement (secondary task performance): Visual accuracy.....	132
3.9.4.2.	Evaluation of validity of LE measurement: Subjective ratings	134
3.9.4.3.	Stability of primary task performance: Auditory accuracy.....	136
3.9.4.4.	Evaluation of viability of LE measurement.....	138
3.10.	Conclusions of Experiment 3.....	141
4	Chapter Four: Introducing adaptive tracking into the refined digit stream dual-task (Experiment 4)	
4.1.	Introduction.....	142
4.2.	Proposed solution.....	142
4.3.	Experiment 4.....	143
4.3.1.	Hypotheses.....	143
4.3.2.	Methods.....	143
4.3.2.1.	Participants.....	143
4.3.2.2.	Procedure.....	144
4.3.3.	Results and Discussion: Descriptive and Inferential statistics.....	148

4.3.3.1.	Secondary task performance: Visual accuracy.....	148
4.3.3.2.	Evaluation of validity of secondary task performance: Subjective ratings	149
4.3.3.3.	Stability of primary task performance: Auditory accuracy.....	150
4.3.3.4.	Evaluation of potentially viability of digit stream dual-task as LE measurement.....	153
4.4.	Conclusions of Experiment 4.....	159

PART FOUR: Digit Stream Dual-Task (Cochlear Implant Simulations & Cochlear Implant Users)

5 Chapter Five: Modifying and using the refined digit stream dual-task in cochlear implant simulations testing (Experiments 5, 6 and 7)

5.1.	Introduction.....	161
5.2.	Proposed solution.....	161
5.3.	Experiment 5.....	163
5.3.1.	Hypotheses.....	163
5.3.2.	Methods.....	163
5.3.2.1.	Participants.....	163
5.3.2.2.	Procedure.....	163
5.3.3.	Results and Discussion: Descriptive statistics.....	169
5.3.3.1.	LE measurement (secondary task performance): Visual accuracy.....	169
5.3.3.2.	Evaluation of validity of LE measurement: Subjective ratings.....	171
5.3.3.3.	Stability of primary task performance: Auditory accuracy.....	172
5.3.3.4.	Evaluation of viability of LE measurement.....	174
5.4.	Conclusions of Experiment 5.....	177
5.5.	Proposed solution.....	177
5.6.	Experiment 6.....	178
5.6.1.	Hypotheses.....	178
5.6.2.	Methods.....	178

5.6.2.1.	Participants.....	178
5.6.2.2.	Procedure.....	178
5.6.3.	Results and Discussion: Descriptive statistics.....	184
5.6.3.1.	LE measurement (secondary task performance): Visual accuracy.....	184
5.6.3.2.	Stability of primary task performance: Auditory accuracy.....	185
5.6.3.3.	Evaluation of viability of LE measurement.....	185
5.7.	Conclusions of Experiment 6.....	191
5.8.	Proposed solution.....	191
5.9.	Experiment 7.....	192
5.9.1.	Hypotheses.....	192
5.9.2.	Methods.....	192
5.9.2.1.	Participants.....	192
5.9.2.2.	Procedure.....	192
5.9.3.	Results and Discussion: Descriptive and Inferential statistics.....	198
5.9.3.1.	LE measurement (secondary task performance): Visual accuracy.....	198
5.9.3.2.	Evaluation of validity of LE measurement: Subjective ratings	200
5.9.3.3.	Stability of primary task performance: Auditory accuracy.....	202
5.9.3.4.	Evaluation of viability of LE measurement.....	204
5.10.	Conclusions of Experiment 7.....	209
6	Chapter Six: Testing the refined digit stream dual-task with cochlear implant users (Experiment 8)	
6.1.	Introduction.....	210
6.2.	Required modifications for CI testing.....	210
6.3.	Experiment 8.....	211
6.3.1.	Hypotheses.....	211
6.3.2.	Methods.....	211
6.3.2.1.	Participants.....	211
6.3.2.2.	Procedure.....	213

6.3.3.	Results and Discussion: Descriptive and Inferential statistics.....	218
6.3.3.1.	LE measurement (secondary task performance): Visual accuracy.....	218
6.3.3.2.	Evaluation of validity of LE measurement: Subjective ratings	220
6.3.3.3.	Stability of primary task performance: Auditory accuracy.....	223
6.3.3.4.	Evaluation of viability of LE measurement.....	225
6.4.	Conclusions of Experiment 8.....	237

PART FIVE: Discussion

7 Chapter Seven: General discussion and future directions

7.1.	Introduction.....	239
7.2.	Summary of all experiments.....	239
7.3.	Summary of key findings.....	240
7.4.	Application of key results to cognitive capacity.....	241
7.5.	Relationship between key results and existing literature.....	245
7.6.	Potential implications of key results.....	246
7.7.	Limitations in current dual-task design.....	249
7.7.1.	CI simulation training.....	249
7.7.2.	Adaptive tracking.....	249
7.7.3.	Task difficulty.....	250
7.7.4.	Primary task stability.....	251
7.7.5.	Perceptual response.....	251
7.7.6.	Scaling LE.....	253
7.7.7.	Subjective ratings.....	254
7.7.8.	Power.....	254
7.8.	Clinical feasibility of LE testing.....	255
7.9.	Conclusion of feasibility of dual-task paradigm.....	257

Appendix	259
-----------------	-----

References	325
-------------------	-----

List of Figures

(Initial number indicates chapter)

1.1.	Schematic of FUEL model.....	25
1.2.	Graphical conceptualisations of FUEL model & time course of effort.....	28
1.3.	Schematic of potential factors involved in listening effort within the CI user	56
1.4.	Schematic of structure and content of thesis.....	77
2.1.	Venn diagrams of resource allocation in dual-task paradigm.....	82
2.2.	Schematic of test conditions for Pilot Study.....	90
2.3.	Schematic of trial structure for Pilot Study.....	91
2.4.	Dot plots of visual recall accuracy & LE (Pilot Study)	97
2.5.	Dot plots of mean task-evoked pupil response (Pilot Study).....	99
2.6.	Dot plots of subjective ratings (Pilot Study).....	102
2.7.	Dot plots of auditory accuracy & primary task stability (Pilot Study).....	104
2.8.	Dot plot of mean SRT (CCRM).....	105
3.1.	Schematic of test conditions for Experiment 1.....	114
3.2.	Schematic of trial structure for Experiment 1.....	115
3.3.	Dot plots of visual accuracy (Experiment 1).....	119
3.4.	Dot plots of auditory accuracy & primary task stability (Experiment 1).....	120
3.5.	Stacked histograms of over-reporting incidence (Experiment 1).....	122
3.6.	Schematic of test conditions for Experiment 2.....	125
3.7.	Schematic of trial structure for Experiment 2.....	125
3.8.	Stacked histograms of over-reporting incidence & dot plot of visual accuracy (Experiment 2).....	128
3.9.	Schematic of test conditions for Experiment 3.....	130

3.10.	Boxplots of visual accuracy & LE (Experiment 3).....	133
3.11.	Boxplots of subjective ratings (Experiment 3).....	134
3.12.	Boxplots of auditory accuracy & primary task stability (Experiment 3).....	137
3.13.	Histograms of the nature of visual response (Experiment 3).....	139
3.14.	Boxplots of over-reporting incidence (Experiment 3).....	140
4.1.	Schematic of test conditions for Experiment 4.....	146
4.2.	Boxplots of visual accuracy (Experiment 4).....	148
4.3.	Boxplots of subjective ratings (Experiment 4).....	149
4.4.	Boxplots of auditory accuracy & primary task stability and dot plot of SRT (Experiment 4).....	152
4.5.	Histograms of the nature of visual response (Experiment 4).....	154
4.6.	Boxplots of over-reporting incidence (Experiment 4).....	155
4.7.	Scatterplot of auditory accuracy versus SRT & exemplars of psychometric function and adaptive track (Experiment 4).....	158
5.1.	Schematic of conditions in CI simulation training (Experiment 5).....	164
5.2.	Schematic of test conditions for Experiment 5.....	167
5.3.	Dot plots of visual accuracy (Experiment 5).....	170
5.4.	Dot plots of subjective ratings (Experiment 5).....	171
5.5.	Dot plots of auditory accuracy & primary task stability and scatterplot of testing parameters (Experiment 5).....	173
5.6.	Scatterplots of auditory accuracy versus channel number & SNR and exemplars of output from channel & noise adaptive tracking (Experiment 5).....	174
5.7.	Stacked histogram of over-reporting incidence & histograms of the nature of visual response (Experiment 5).....	176
5.8.	Schematic of conditions in CI simulation training (Experiment 6).....	179
5.9.	Schematic of test conditions for Experiment 6.....	182
5.10.	Dot plots of visual accuracy & LE (Experiment 6).....	184

5.11.	Dot plots of auditory accuracy & primary task stability and scatterplot of testing parameters (Experiment 6).....	186
5.12.	Stacked histograms of over-reporting incidence (Experiment 6).....	187
5.13.	Histograms of the nature of visual response (Experiment 6).....	188
5.14.	Scatterplots of auditory accuracy versus channel number and SNR and exemplars of output from channel & noise adaptive tracking (Experiment 6).....	190
5.15.	Schematic of conditions in CI simulation training (Experiment 7).....	193
5.16.	Schematic of test conditions for Experiment 5.....	195
5.17.	Boxplots of visual accuracy (Experiment 7).....	199
5.18.	Boxplots of subjective ratings (Experiment 7).....	200
5.19.	Boxplots of auditory accuracy & primary task stability and scatterplot of testing parameters (Experiment 7).....	203
5.20.	Histograms of the nature of visual response (Experiment 7).....	205
5.21.	Boxplots of over-reporting incidence (Experiment 7).....	206
5.22.	Scatterplots of auditory accuracy versus channel number and SNR, boxplots of auditory accuracy in training and testing sessions, exemplars of output from channel adaptive tracking (Experiment 7).....	208
6.1.	Pie charts of demographics of CI users (Experiment 8).....	212
6.2.	Schematic of test conditions for Experiment 8.....	215
6.3.	Boxplots of visual accuracy & LE (Experiment 8).....	219
6.4.	Boxplots of subjective ratings (Experiment 8).....	221
6.5.	Boxplots of auditory accuracy & primary task stability (Experiment 8).....	224
6.6.	Dot plot of auditory accuracy in quiet and exemplars of output from noise adaptive tracking (Experiment 8).....	227
6.7.	Boxplots of subjective ratings parsed into “successes” & “failures” (Experiment 8).....	229
6.8.	Boxplots of visual accuracy & LE parsed into “successes” & “failures” (Experiment 8).....	230

6.9.	Boxplots of auditory accuracy & primary task stability parsed into “successes” & “failures” (Experiment 8).....	231
6.10.	Scatterplots of actual versus intended auditory performance in noise (Experiment 8).....	232
6.11.	Histograms of the nature of visual response (Experiment 8).....	234
6.12.	Boxplots & stacked histograms of over-reporting incidence parsed into “successes” & “failures” (Experiment 8).....	235
6.13.	Boxplots of over-reporting incidence for entire CI cohort (Experiment 8).....	236
7.1.	Venn diagrams of theorised resource allocation for dual-tasking in quiet versus visual task alone.....	242
7.2.	Venn diagrams of theorised resource allocation for dual-tasking in quiet versus visual task alone in the case of CI simulations.....	242
7.3.	Venn diagrams of theorised resource allocation for dual-tasking in quiet & in noise (CI simulations).....	243
7.4.	Venn diagrams of theorised resource allocation for dual-tasking in quiet in CI simulations versus actual CI users.....	243
7.5.	Venn diagrams of theorised resource allocation for dual-tasking in quiet & in noise in CI simulations versus actual CI users.....	244

List of Tables

(Initial number indicates chapter)

2.1.	Demographic detail of CI users: aetiology and type of deafness & type and length of CI experience (Pilot Study)	84
2.2.	Demographic detail of CI users: age, academic achievement, language preference & knowledge of sign language (Pilot Study)	84
2.3.	Protocol for Pilot Study.....	95
3.1.	Protocol for Experiment 1.....	117
3.2.	Protocol for Experiment 2.....	126
3.3.	Protocol for Experiment 3.....	131
3.4.	Mean and SE of visual accuracy (Experiment 3).....	133
3.5.	Mean and SE of subjective ratings (Experiment 3).....	135
3.6.	Mean and SE of auditory accuracy (Experiment 3).....	137
4.1.	Protocol for Experiment 4.....	147
4.2.	Mean and SE of subjective ratings (Experiment 4).....	150
5.1.	Protocol for Experiment 5.....	168
5.2.	Protocol for training (Experiment 6).....	180
5.3.	Protocol for testing (Experiment 6).....	183
5.4.	Protocol for training (Experiment 7).....	194
5.5.	Protocol for testing (Experiment 7).....	196
5.6.	Mean and SE of visual accuracy (Experiment 7).....	199
5.7.	Mean and SE of subjective ratings (Experiment 7).....	201
6.1.	Protocol for Experiment 8.....	217
6.2.	Mean and SE of visual accuracy (Experiment 8).....	219
6.3.	Mean and SE of subjective ratings (Experiment 8).....	221
6.4.	<i>p</i> -values for statistically significant pairwise comparisons of subjective ratings across test conditions (Experiment 8).....	222
6.5.	<i>p</i> -values for statistically insignificant pairwise comparisons of subjective ratings across test conditions (Experiment 8).....	222

List of Appendices

(Initial number indicates chapter)

1.1.	Protocol guidelines for a CI audiological assessment.....	259
2.1.	Protocol for additional assessments (Pilot Study)	260
2.2.	Methodological detail of CANTAB (Pilot Study).....	261
2.3.	Dot plots of performance in sub-tests of WASI and Reading Span.....	262
2.4.	Dot plots of performance in sub-tests of CANTAB Span.....	263
2.5.	Boxplots of variability in resting pupil size.....	264
2.6.	Boxplots of variability in task-evoked pupil response.....	265
3.1.	Explanation regarding choice of visual accuracy scoring method.....	266
4.1.	Psychometric functions and adaptive tracks for NH controls 1-4 (Experiment 4: noise tracking).....	270
4.2.	Psychometric functions and adaptive tracks for NH controls 5-8 (Experiment 4: noise tracking).....	271
4.3.	Psychometric functions and adaptive tracks for NH controls 9-12 (Experiment 4: noise tracking).....	272
4.4.	Psychometric functions and adaptive tracks for NH controls 13-16 (Experiment 4: noise tracking).....	273
4.5.	Psychometric functions and adaptive tracks for NH controls 17-20 (Experiment 4: noise tracking).....	274
4.6.	Psychometric functions and adaptive tracks for NH controls 21-24 (Experiment 4: noise tracking).....	275
4.7.	Psychometric functions and adaptive tracks for NH controls 25-28 (Experiment 4: noise tracking).....	276
4.8.	Psychometric functions and adaptive tracks for NH controls 29-30 (Experiment 4: noise tracking).....	277

5.1.	Psychometric functions and adaptive tracks for NH controls 1-3 (Experiment 5: noise tracking).....	278
5.2.	Psychometric functions and adaptive tracks for NH controls 4-6 (Experiment 5: noise tracking).....	279
5.3.	Psychometric functions and adaptive tracks for NH controls 1-3 (Experiment 5: channel tracking).....	280
5.4.	Psychometric functions and adaptive tracks for NH controls 4-6 (Experiment 5: channel tracking).....	281
5.5.	Psychometric functions and adaptive tracks for NH controls 1 & 2 (Experiment 6: noise tracking).....	282
5.6.	Psychometric functions and adaptive tracks for NH controls 3 & 4 (Experiment 6: noise tracking).....	283
5.7.	Psychometric functions and adaptive tracks for NH controls 5 & 6 (Experiment 6: noise tracking).....	284
5.8.	Psychometric functions and adaptive tracks for NH controls 7 & 8 (Experiment 6: noise tracking).....	285
5.9.	Psychometric functions and adaptive tracks for NH controls 9 & 10 (Experiment 6: noise tracking).....	286
5.10.	Psychometric functions and adaptive tracks for NH controls 1 & 2 (Experiment 6: channel tracking).....	287
5.11.	Psychometric functions and adaptive tracks for NH controls 3 & 4 (Experiment 6: channel tracking).....	288
5.12.	Psychometric functions and adaptive tracks for NH controls 5 & 6 (Experiment 6: channel tracking).....	289
5.13.	Psychometric functions and adaptive tracks for NH controls 7 & 8 (Experiment 6: channel tracking).....	290
5.14.	Psychometric functions and adaptive tracks for NH controls 9 & 10 (Experiment 6: channel tracking).....	291
5.15.	Psychometric functions and adaptive tracks for NH controls 1 & 2 (Experiment 7: channel tracking).....	292
5.16.	Psychometric functions and adaptive tracks for NH controls 3 & 4 (Experiment 7: channel tracking).....	293

5.17.	Psychometric functions and adaptive tracks for NH controls 5 & 6 (Experiment 7: channel tracking).....	294
5.18.	Psychometric functions and adaptive tracks for NH controls 7 & 8 (Experiment 7: channel tracking).....	295
5.19.	Psychometric functions and adaptive tracks for NH controls 9 & 10 (Experiment 7: channel tracking).....	296
5.20.	Psychometric functions and adaptive tracks for NH controls 11 & 12 (Experiment 7: channel tracking).....	297
5.21.	Psychometric functions and adaptive tracks for NH controls 13 & 14 (Experiment 7: channel tracking).....	298
5.22.	Psychometric functions and adaptive tracks for NH controls 15 & 16 (Experiment 7: channel tracking).....	299
5.23.	Psychometric functions and adaptive tracks for NH controls 17 & 18 (Experiment 7: channel tracking).....	300
5.24.	Psychometric functions and adaptive tracks for NH controls 19 & 20 (Experiment 7: channel tracking).....	301
5.25.	Psychometric functions and adaptive tracks for NH controls 21 & 22 (Experiment 7: channel tracking).....	302
5.26.	Psychometric functions and adaptive tracks for NH controls 23 & 24 (Experiment 7: channel tracking).....	303
5.27.	Psychometric functions and adaptive tracks for NH controls 25 & 26 (Experiment 7: channel tracking).....	304
5.28.	Psychometric functions and adaptive tracks for NH controls 27 & 28 (Experiment 7: channel tracking).....	305
5.29.	Psychometric functions and adaptive tracks for NH controls 29 & 30 (Experiment 7: channel tracking).....	306
5.30.	Breakdown of number of participants requiring each number of channels and noise level for vocoding (Experiment 7).....	307
6.1.	Visual accuracy of two test runs across test conditions (Experiment 8).....	308
6.2.	Subjective ratings of two test runs across test conditions (Experiment 8).....	309

6.3.	Auditory accuracy of two test runs across test conditions (Experiment 8).....	310
6.4.	Psychometric functions and adaptive tracks for CI users 1 & 2 (Experiment 8: noise tracking).....	311
6.5.	Psychometric functions and adaptive tracks for CI users 3 & 4 (Experiment 8: noise tracking).....	312
6.6.	Psychometric functions and adaptive tracks for CI users 5 & 6 (Experiment 8: noise tracking).....	313
6.7.	Psychometric functions and adaptive tracks for CI users 7 & 8 (Experiment 8: noise tracking).....	314
6.8.	Psychometric functions and adaptive tracks for CI users 9 & 10 (Experiment 8: noise tracking).....	315
6.9.	Psychometric functions and adaptive tracks for CI users 11 & 12 (Experiment 8: noise tracking).....	316
6.10.	Psychometric functions and adaptive tracks for CI users 13 & 14 (Experiment 8: noise tracking).....	317
6.11.	Psychometric functions and adaptive tracks for CI users 15 & 16 (Experiment 8: noise tracking).....	318
6.12.	Psychometric functions and adaptive tracks for CI users 17 & 18 (Experiment 8: noise tracking).....	319
6.13.	Psychometric functions and adaptive tracks for CI users 19 & 20 (Experiment 8: noise tracking).....	320
6.14.	Psychometric functions and adaptive tracks for CI users 21 & 22 (Experiment 8: noise tracking).....	321
6.15.	Psychometric functions and adaptive tracks for CI users 23 & 24 (Experiment 8: noise tracking).....	322
6.16.	Psychometric functions and adaptive tracks for CI user 25 (Experiment 8: noise tracking).....	323

PART ONE:

INTRODUCTION

CHAPTER ONE:

Literature Review

1.1. Overview of the context

Deafness is one of the most common disabilities, with approximately 360 million people afflicted worldwide, i.e. 5% of the global population (WHO, 2017; Winn et al., 2015). Currently, at least 700,000 of these are recipients of a neuroprosthesis called the cochlear implant, an electronic device which bypasses the damaged inner ear and directly stimulates the auditory nerve to provide a sensation of sound (Cochlear, 2017), even in cases of complete obliteration of sensory cells (Kral et al., 2016). The cochlear implant (CI) is argued to be unprecedented in its triumph as biomedical technology, with profoundly and even totally deaf individuals achieving remarkable results in everyday listening, including high accuracy of speech understanding and intelligibility in speech production (Lazard et al., 2012a, b; Kral et al., 2016; Peterson et al., 2010).

The established efficacy of the CI, combined with ever-increasing economic viability in technology provision, means that the popularity of cochlear implantation as the chosen route for hearing intervention is growing (Pisoni et al., 1999; Technavio, 2016; Peterson et al., 2010). Using the current growth rates of units being shipped annually, it is projected that 96,000 CIs will be sold and globally distributed within the year of 2020 alone (Technavio, 2016). With an estimated 134 million children being born each year, of which one per thousand will have a hearing loss that can only be ameliorated with a CI, the call for cochlear implantation is unlikely to disappear (Hochmair, 2013). In fact, the demand for CIs is more likely to burgeon even further due to the relaxation in criteria for CI candidacy to now include hearing impaired adults with some residual hearing and even the elderly population with age-related hearing loss (Blamey et al., 2013; Lamb & Archbold, 2013; Worrall & Hickson, 2003; Yang & Cosetti, 2016).

Accompanying this enhanced demand are also heightened expectations regarding the CI's ability to function (more or less) as well as an undamaged auditory system (Kral et al., 2016; Peterson et al., 2010). However, the notable achievements of CI technology are marred by the presence of substantial variability in post-operative outcomes: while some CI users succeed to such a level that they can even manage telephone conversations, others never develop useable speech and language skills (Cohen et al., 1999; Niparko et al., 2010; Dorman et al., 2008; Geers et al., 2011a, b; Tobey et al., 2013). Indeed, in the case of disyllabic word recognition one year after cochlear implantation, the maximum possible range of 0% and 100% is found (with the median being 70%: Lazard et al., 2010).

Alarming, when poor attainment occurs, it can continue to persist even with years of CI usage and experience, or when listening conditions are optimal, i.e. in quiet (Blamey et al., 2013; Anderson et al., 2017a, b; Holden et al., 2013; Lazard et al., 2010). In fact, the variability in CI outcome occurs so consistently (with multiple CI centres worldwide reporting it in both congenitally deafened children and postlingually deafened adults) that it has become a hallmark of the technology (Peterson et al. 2010; Lee et al., 2005; Kral et al., 2016). To further exacerbate matters, only 20% of this variance in CI outcomes can reliably be accounted for (within multifactor models) by peripheral issues such as comorbidity, demographics, device, CI surgery, and/or hearing history (Lazard et al., 2012a; O'Donoghue et al., 2000; Blamey et al., 2013; Green et al., 2007; Holden et al., 2013).

The inexplicability of the remaining 80% of the variance, combined with booming cochlear implantation rates, means that this inconsistency in functional outcome is a pressing clinical problem, especially in cases of successful CI surgery (Kral et al., 2016). It could even hinder, or stagnate, technological innovation and development of rehabilitation techniques (O'Donoghue et al., 2016; Lazard et al., 2012a, b). However, with recent neuroscientific advances in tools to interrogate neuronal function, the nature and extent of the impact of auditory deprivation on the human central nervous system (and not just the periphery) can be further elucidated beyond the previous physiological understanding derived from animal models (O'Donoghue et al., 2016; Lyness et al., 2013; Wijayasari et al., 2017). It is now becoming increasingly apparent that neurocognitive factors are just as, if not more, important than the peripheral factors typically considered necessary for succeeding with the CI (Lazard et al., 2012b; Moore & Shannon, 2009; Kral et al., 2016; O'Donoghue et al., 2016).

A corollary of this focus on the neurocognitive aspect of hearing rehabilitation is the recognition of the actual cognitive burden of deafness (and the CI) on the brain (Pichora-Fuller et al., 2016; McGarrigle et al., 2014). In particular, there is escalating interest in the construct termed “listening effort” (Pichora-Fuller et al., 2016; McGarrigle et al., 2014). What has not yet been explored is whether listening effort contributes to the variance in CI outcomes, even in the most proficient of CI users. Indeed, it is now being argued that, if the neurocognitive interactions between the brain and the CI are investigated (and the impact of speech processing demand and the true cognitive toll imposed by the CI ascertained), then the potential for the rehabilitation of the CI user and the CI technology can be fully realised (Kral et al., 2016; O'Donoghue et al., 2016). Therefore, recognition and clinical measurement of listening effort could provide a valuable insight into the degree of this cognitive toll being experienced by CI users.

1.2. Understanding listening effort: a literature review

The surge of interest in listening effort does not automatically lead to an understanding of the relevance of listening effort to the CI population. Thus, in order to provide a rationale for this thesis, this literature review will explore and attempt to answer the following questions:

- What is listening effort?
- Why is listening effort important?
- How can listening effort be measured?

1.2.1. What is listening effort?

Listening effort (LE), as a construct, has been a somewhat elusive concept (McGarrigle et al., 2014; Pichora-Fuller et al., 2016). Frequently, LE has been referred to as being the attention and cognitive resources required to understand speech (Hicks & Tharpe, 2002; Gosselin & Gagné, 2010; Fraser et al., 2010; Picou et al., 2011). However, there has been considerable disagreement about whether this is indeed the correct way to conceptualise LE (McGarrigle et al., 2014). For instance, it has been disputed whether LE actually is exclusive to speech processing, or whether it applies to other types of auditory stimuli too, such as music or environmental sounds (McGarrigle et al., 2014). After all, auditory perception is a multidimensional phenomenon enabling auditory object formation and perception of a complex auditory scene (Shinn-Cunningham & Best, 2008; Bizley & Cohen, 2013). Within this auditory scene, speech processing is only a single (though important) element (McGarrigle et al., 2014).

In response, there has been a proliferation of working definitions for LE, with no particularly robust candidate to rely upon (McGarrigle et al., 2014; Lemke & Besser, 2016). This has been partly blamed on the lack of strong theoretical underpinnings due to the relative immaturity of this research field (McGarrigle et al., 2014). Indeed, there has even been debate as to whether LE is a valid construct at all, with it being posited that the idea of LE is simply a manifestation of this paucity in knowledge (McGarrigle et al., 2014).

Fortunately, however, with the importance of LE becoming increasingly acknowledged by the audiological profession, the Eriksholm Workshop on Hearing Impairment and Cognitive Energy specifically developed a heuristic theoretical model called the Framework for Understanding Effortful Listening, or FUEL (Figure 1.1: Pichora-Fuller & Kramer, 2016).

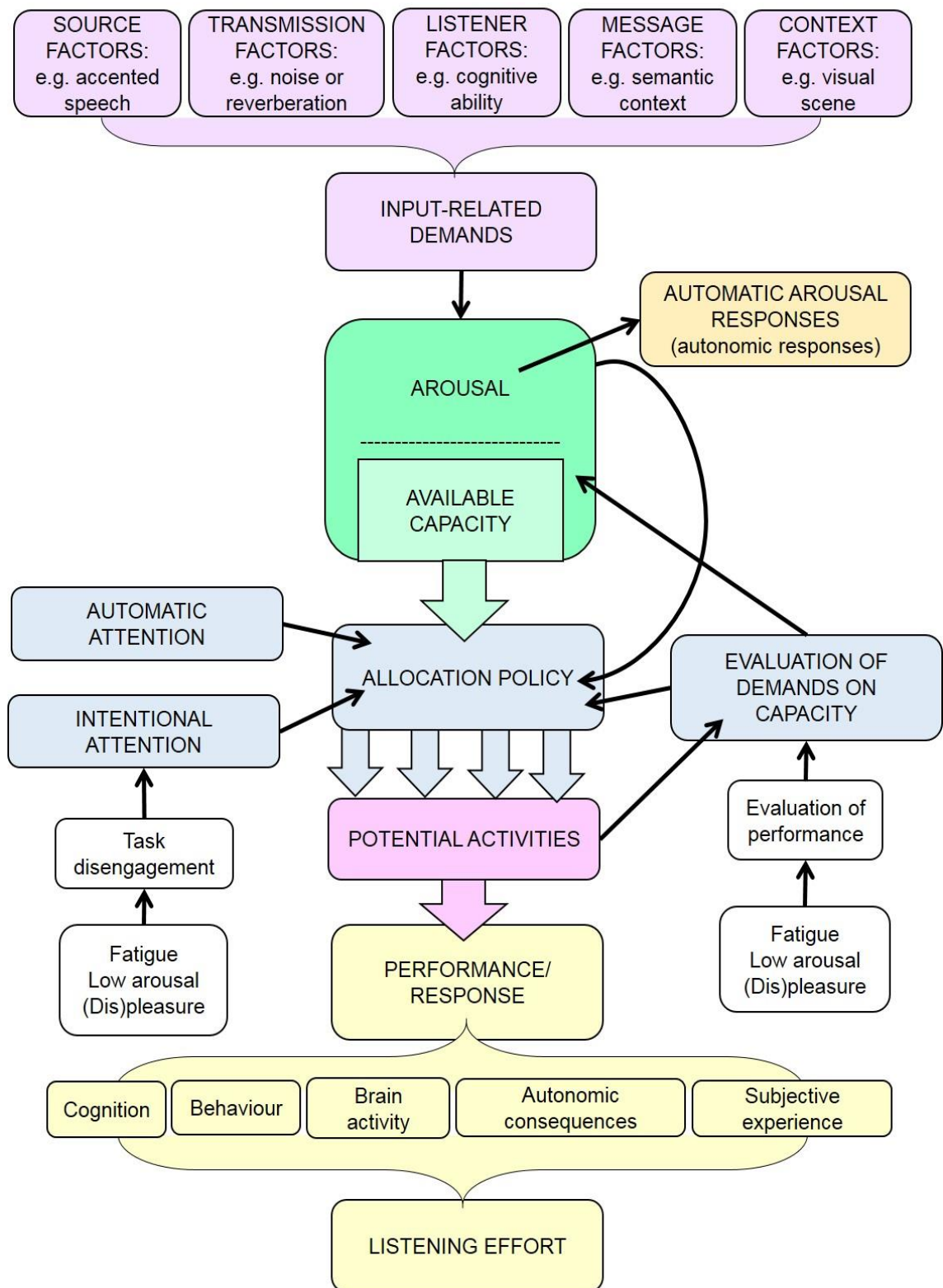


Figure 1.1: Schematic of the theoretical principles underlying the model of the Framework for Understanding Effortful Listening (i.e. FUEL). Modified from Pichora-Fuller et al. (2016) and Kahneman (1973).

1.2. Understanding listening effort

The core of FUEL is the seminal Capacity Model of Attention, which posits that, fundamentally, the brain can provide a finite pool of cognitive resources, i.e. the available capacity (Kahneman, 1973). This pool of cognitive resources fluctuates in accordance with the allocation policy, which is defined to be the executive function that determines how the cognitive resources are distributed for the specific task execution at the time (Kahneman, 1973; Rudner, 2016). How the cognitive resources are distributed is dependent upon the input-related demands (which involves a wide assortment of factors, such as whether or not listening becomes more challenging due to the presence of noise or reverberation), as well as the competing tasks also vying for priority (Pichora-Fuller et al., 2016; Wingfield, 2016; Watson & Strayer, 2010; Edwards, 2016). The allocation policy ultimately yields the performance, together with all the associated responses, such as the behaviour, brain activity, and/or the autonomic consequences entailed in the actual task execution (Kahneman, 1973).

According to the tenets of the Capacity Model, the allocation policy described by FUEL is itself modulated by four factors: first of all, the internal dispositions which influence the action of involuntary attention (i.e. automatic attention, the intrinsic tendency to orient attention towards the sudden presence of a novel stimulus); secondly, momentary intentions, such as the following of specific instructions (i.e. intentional attention); thirdly, the evaluation of the demands being placed on capacity; and finally, the effects of arousal (Kahneman, 1973; Pichora-Fuller et al., 2016).

Having decided on the underlying neurocognitive architecture, FUEL then considered the reality of everyday listening and how listening can become effortful, whereby the concept of adversity became key (Pichora-Fuller et al., 2016; Lemke & Besser, 2016). This is because most listening situations are rarely ideal, with a wide variety possible in the nature of adverse conditions likely to be encountered: i.e. impoverishment within the acoustic signal, degraded transmission of the sound, or unclear communication (Lemke & Besser, 2016; Denes & Pinson, 1963; Humes & Bess, 2013; Mattys et al., 2012).

Adversity is thus conceptualised to be the mismatch between external demand and the internal resources available to meet these demands (Lemke & Besser, 2016).

Accordingly, LE is defined as the deliberate allocation of mental resources (or cognitive capacity) to overcome obstacles in goal pursuit when carrying out a task that involves listening (Pichora-Fuller et al., 2016; Lemke & Besser, 2016).

The adversity encountered in the input-related demands is a notion highly familiar with audiologists and is the frequent target for battle in hearing rehabilitation, particularly with the use of assistive technology (Pichora-Fuller, 2016; Erber, 1988).

1.2. Understanding listening effort

However, since LE is the deliberate allocation of cognitive capacity to overcome demands, it is not just these input-related demands that are important; the individual's decision regarding when and to what extent they will expend effort in that particular goal pursuit is also crucial (Pichora-Fuller, 2016). It cannot be assumed that capacity will automatically be allocated whenever demand increases, or that the individual has sufficient capacity to meet these demands (Pichora-Fuller, 2016; Lemke & Besser, 2016; Matthen, 2016; Richter, 2016). The individual's actual willingness to devote cognitive capacity implicates a multitude of factors, which predominantly involve the evaluation of demands and also the intentional attention within the FUEL model (Pichora-Fuller, 2016).

Evaluation of demand is key to determining the nature of the allocation policy that is, in turn, determined by the input-related demands (Pichora-Fuller et al., 2016). FUEL already posits low arousal, fatigue and displeasure to reduce the likelihood of resource allocation (Pichora-Fuller et al., 2016). Evaluation of demand may also involve the consideration of competing demands on capacity by multiple possible activities (Eckert et al., 2016; Pichora-Fuller et al., 2016). Intentional attention introduces the dimension of motivation into the individual's decision to devote capacity (Pichora-Fuller, 2016). Motivation involves the likelihood of even engaging in the task in the first place, as well as the sustaining of execution, which then introduces factors such as whether or not the listening goal has sufficient value to the listener (such as the reward obtained, be it personal, social or related to pleasure) or whether successful performance is important to the listener (Richter, 2016; Matthen, 2016; Pichora-Fuller, 2016; Pichora-Fuller et al., 1998; Ryan et al., 1986, 1995; Chasteen et al., 2015).

Thus, to address this additional element of motivation, FUEL also includes a three-dimensional graphical conceptualisation, in an attempt to illustrate how the effort expended is moderated by motivational intensity, as well as how allocation policy has been influenced and implemented according to the demand encountered (Pichora-Fuller et al., 2016: Figure 1.2a). Superimposed on this three-dimensional plot is yet another plane: that of time. This is in an attempt to show how effort exerted by the individual might change over the course of time for a given activity (Pichora-Fuller et al., 2016).

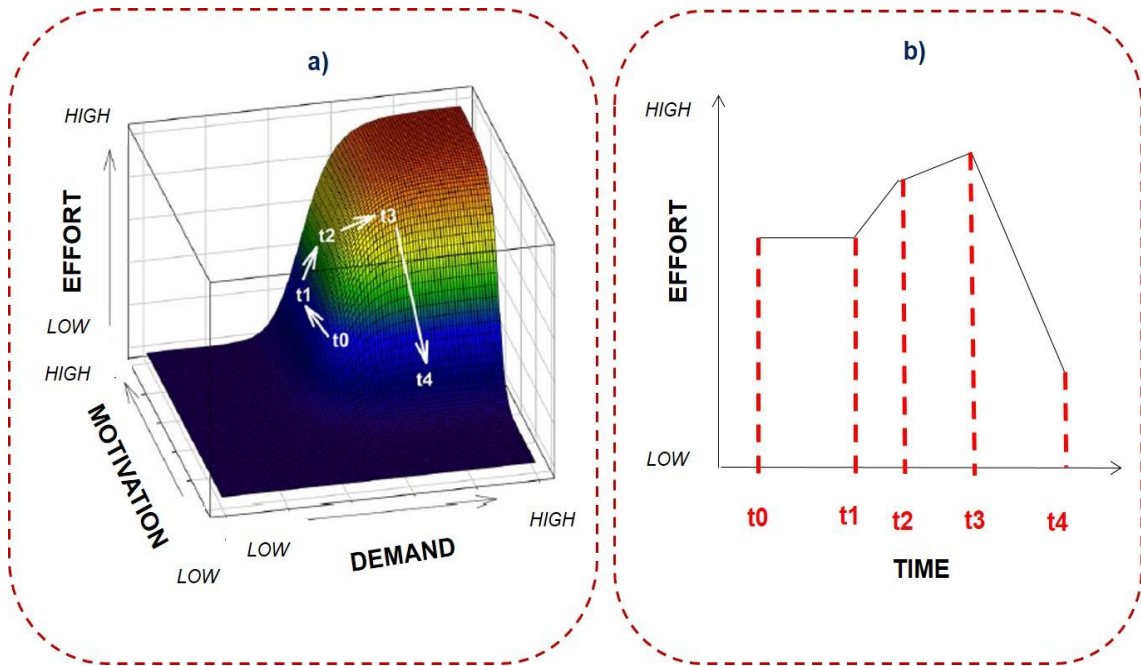


Figure 1.2: Plots conveying the relationship of arousal; motivation; and time with the effort exerted (according to the FUEL model). **(a)** The novel three-dimensional conceptualisation given by Pichora-Fuller et al. to complement the FUEL model. The three axes depicting the interactions of motivation of the individual and demand of task on the effort exerted; superimposed on this plot is also a timeline (i.e. t_0 , t_1 , t_2 , t_3 and t_4) to indicate how all three axes could fluctuate over the course of time. **(b)** Elucidation of the meaning and potential influence of time on the effort exerted (this time course corresponds to the timeline t_0 - t_4 indicated in panel **a**). Modified from Pichora-Fuller et al. (2016).

To make more explicit precisely what the temporal effect could entail on the interactions involved in effort exertion, another graphical representation is provided by FUEL (Pichora-Fuller et al., 2016: Figure 1.2b). Here, effort is described as being the composite of both the input-related demands and also motivational intensity, and is given as a function of time. A hypothetical scenario is provided to explore how effort could change over time: initially, there is little change in effort, as well as demand staying constant (at a low level), but motivation becomes enhanced as the individual's engagement in the task increases (t_0 - t_1); motivation then remains level, but demand may subsequently increase with the introduction of an adverse listening condition (e.g. background noise), meaning effort level has to ramp up to sustain performance (t_1 - t_2 - t_3); the input-related demand then stays constant (with the adverse listening condition having become more established), but the individual's motivation declines as the task at hand, potentially due to listening-related fatigue, leading to a reduction in effort to accompany task disengagement (t_3 - t_4) (Pichora-Fuller et al., 2016).

1.2. Understanding listening effort

The bell-shaped nature of the effort-motivation-demand relationship described by FUEL was inspired by the Hebbian account of Yerkes-Dodson's law, which describes the empirical relationship between arousal level and performance achieved (Pichora-Fuller et al., 2016; Yerkes & Dodson, 1908; Hebb, 1955; Diamond et al., 2007; Hopstaken et al., 2015). This is because the construct of arousal itself influences the individual's assessment of the importance of success in task accomplishment, as well as the value of expending cognitive resources to meet the task's demand on cognitive capacity.

Within FUEL, arousal itself is posited to be controlled by two factors: the input-related demands (and their effect on both brain and body as afflicted by task performance); as well as a miscellany of "determinants", such as the intensity of a given stimulation or the physiological effects of "drive" states (Kahneman, 1973). The vagueness affiliated with the original use of the term "determinants" reflects the fact that arousal itself is of such manifold complexity that it has been particularly difficult to explain precisely, especially in terms of physiological mechanisms (Aston-Jones & Cohen, 2005). Yet its importance cannot be denied, and it is known to be closely related to other phenomena such as sleep, attention, anxiety, stress, as well as motivation itself (Pichora-Fuller et al., 2016). Indeed, with regards to attention, arousal has been cited to be the cost of resisting distraction, with bolstered motor tension as well as other types of autonomic manifestations (Kahneman, 1973). Dampened arousal results in drowsiness and (ultimately) sleep, whereas heightened arousal can be a significant facilitator for behaviours. However, if arousal levels become excessive, distractibility and anxiety become the (unwanted) consequences (Pichora-Fuller et al., 2016).

Because of the role of time involved in effort, fatigue too is of importance when it comes to level of effort being expended during the course of task execution and the elicited subjective experience of this exertion (Pichora-Fuller et al., 2016). Fatigue is multifaceted, with all kinds of manifestations possible, such as physical fatigue; emotional or affective fatigue; as well as mental or cognitive fatigue (Hornsby et al., 2016). Within the context of FUEL, the sustained task performance engenders listening fatigue (a subclass of mental or cognitive fatigue) due to the continued application of mental effort draining the finite capacity of cognitive resources (Kahneman, 1973; Hornsby et al., 2016). This then brings about the experience of weariness or tiredness, as well as reduced vigour, vitality and energy (Hornsby et al., 2016; Pichora-Fuller et al., 2016). More importantly, this fatigue leads to increased likelihood of task disengagement, due to decreasing motivation leading to self-perceived (and also objectively measurable) difficulties in concentration, attention, clear thinking or memory (Hetú et al., 1988; Kramer et al., 2006).

1.2. Understanding listening effort

In fact, it is argued that this subjective fatigue actually serves an adaptive and goal-directed function by forcing the individual to evaluate current behaviour (i.e. the cognitive capacity being allocated) to achieve a reward from completion, or continuation, of a particular task (Hockey, 2013; Hornsby et al., 2016; Eckert et al., 2016). This adaptive control is conducted in order to optimise behaviour and ensure cognitive capacity is not being consumed unnecessarily, or inefficiently (Eckert et al., 2016). According to this argument, a type of neuroeconomics is essentially performed, with a cost-benefit analysis ascertaining the nature of the effort-reward relationship (Matthen, 2016; Hockey, 2013; Eckert et al., 2016). Should the effort-reward relationship prove to be unfavourable, motivation in task continuation may diminish (Matthen, 2016; Eckert et al., 2016; Hockey, 2013). However, if sufficient reward is achieved, the task has value (which can be personal or social in nature), which can associate sensations of pleasure with the reward output (Matthen, 2016; Ryan et al., 1986, 1995; Chasteen et al., 2015).

The elements of motivation and also physiological state (i.e. arousal and fatigue) thus implicate the individual's personal state (Lemke & Besser, 2016). Therein lies another realm of complexity, because the individual's personal state involves psychosocial ecology too (Pichora-Fuller, 2016; Engel, 1977; Frankish et al., 1996; Stineman & Streim, 2010). The psychosocial factors relevant to FUEL involve stress, social support, self-efficacy, and also stigma, with all able to modulate both the individual's evaluation of demand and also the individual's intentional attention (Pichora-Fuller, 2016). In particular, these factors possess the potential to influence both the listener's appraisal of the actual level of demand being imposed by the task and also their ability and capacity to meet those demands, with both immediate and long-term consequences of social participation being considered (Pichora-Fuller, 2016).

For instance, increased level of stress and a low sense of self-efficacy (the individual's belief in their own ability and capacity to meet demand given attainment) would hinder the individual's willingness to allocate resources (Pichora-Fuller, 2016; Bandura, 1997). On the other hand, social support provided by significant others (e.g. financial aid, advice, empathy, or constructive feedback) might mitigate any poor sense of self-efficacy, or stigma (the identification of others or self as having a characteristic that is devalued in a social context), by promoting the use of effective coping strategies and counteracting any negative interpretation of adverse events (Southall et al., 2010; Broadhead et al., 1983, 1988; Cohen, 2004; Cohen & Wills, 1985).

1.2. Understanding listening effort

Because of this involvement of psychosocial ecology, perceived effort may even become unrelated to the actual allocation of cognitive resources (Pichora-Fuller, 2016; Lemke & Besser, 2016). Indeed, perceived effort could arise at any time (even during processing that is close to being effortless), because of the multidimensional nature of the appraisal involved and how it determines the level at which input-related demands becomes adverse for the individual (Lemke & Besser, 2016).

Thus, in summary, LE is ultimately the culmination of a vast assembly of influences and their dynamics with each other. This involves the nature of adversity and how it is appraised. This, in turn, is affected by factors such as arousal, fatigue, personal state and psychosocial ecology. In addition, motivation and the cost-benefit analysis of effort versus reward each have a role. This means that LE also includes the evaluation of demand and intentional attention and how they determine allocation policy. So, despite the concise definition of LE being the deliberate allocation of mental resources (or cognitive capacity) to overcome obstacles in goal pursuit when carrying out a task that involves listening, the actual construct of LE itself is by no means straightforward.

However, in spite of the inherent complexity of LE, a certain level of reassurance can be taken from the fact that the FUEL model is based on well-established principles of cognition and hearing research (Pichora-Fuller et al., 2016; Edwards, 2016).

Furthermore, the general concept of resources, with regards to the brain exerting itself, is a century-old notion with original references citing psychic energy and the perceptual consequences that arise due to its differential allocation (Titchener, 1908). In addition, care was taken within FUEL to incorporate the actuality of the complex world of sound (such as the integration of principles of Auditory Scene Analysis), as well as the interface of language processing (incorporating the Ease of Language Understanding model) (Pichora-Fuller et al., 2016; Edwards, 2016; Bregman, 1990; Rönnberg et al., 2008).

Thus, there is a robust foundation of research literature supporting FUEL. Not only this, there is a comforting level of consistency between the theory of FUEL and the underlying known neurobiology of cognition, especially regarding attention and memory (Eckert et al., 2016). Within the realm of speech recognition research, neuroimaging studies have consistently demonstrated that challenging listening conditions appear to elicit upregulation of the cingulo-opercular system (i.e. the bilateral dorsal cingulate; inferior frontal; and anterior insula regions) and also upregulation of the fronto-parietal systems (i.e. the precentral sulcus; dorsolateral prefrontal cortex; intraparietal sulcus; and inferior parietal lobule) (Dosenbach et al., 2008; Eckert et al., 2016).

Also, promisingly, there are countless studies that have associated these systems with attention and intention, particularly within the function of monitoring and optimising performance related to adaptive control (Eckert et al., 2016). It has been shown that not only does recruitment of these systems appear to lend to improvement in performance, but also that sustained engagement of these systems is correlated with the subjective experience of effort and fatigue (e.g. Vaden et al., 2013; Luks et al., 2007; Menon & Uddin, 2010; Cole et al., 2013; Walsh et al., 2011; Eckert et al., 2016). The notion of a cost-benefit analysis occurring within allocation policy is also connected with these two networks, with the adaptive control also appearing to take into consideration the relative value of the listening task at hand, with the neural systems being utilised more when the value of listening outweighs the cognitive cost necessitated (Kouneiher et al., 2009; Paulus et al., 2003; Eckert et al., 2016)

Therefore, reasonable confidence can be presumed with the application of FUEL and its principles regarding the existence and components of LE.

1.2.2. Why is listening effort important?

Defining and elucidating LE alone is insufficient in accounting for the importance and relevance of LE for the hearing impaired population and for cochlear implant (CI) users. Therefore, there is the need to revisit and review the reality and consequences of hearing impairment (or loss) on the individual.

For those with intact hearing, it could be argued that everyday listening is generally performed virtually effortlessly, and even perceived to be relatively easy (at least in good acoustic conditions), i.e. minimal LE is induced to produce satisfactory performance (McGarrigle et al., 2014). However, once hearing impairment is present, there is immediately diminished fidelity in the sensory information that the brain has to contend with (Mattys et al., 2012).

This will inevitably lead to greater cognitive capacity being deployed in an attempt to enable comprehension and memorisation of, and appropriate responding to, the perceived auditory object or event (Mattys & Wiget, 2011; Rönnberg et al., 2013). Indeed, one of the most common complaints plaguing the deaf and hard-of-hearing population is that listening is taxing, even to the point of being too hard (McGarrigle et al., 2014; Pichora-Fuller et al., 2016; Gosselin & Gagné, 2010).

Linked to this complaint is also a cluster of detrimental socioeconomic consequences (Winn et al., 2015). For instance, occupational performance and productivity within the deaf population both appear to be especially vulnerable to LE levels (McGarrigle et al., 2014; Winn et al., 2015). Indeed, significantly higher incidence of sick leave is reported for deaf workers compared to their normal hearing counterparts (Kramer et al., 2006), with LE-induced fatigue and mental distress included in the cited reasons (Kramer et al., 2006; McGarrigle et al., 2014; Hornsby, 2013). In addition, a positive association has been yielded between poor hearing thresholds and the need for extended recovery time at the end of the working day (Nachtegaal et al., 2009). Difficulties in coping with hearing loss have also been cited to lead to the increased probability of early retirement, and even attainment of employment in the first place, with higher rates of unemployment for young adults who are deaf (Danemark & Gellerstedt, 2004; Parving & Christensen, 1993; Järvelin et al., 1997). It is possible that LE is a contributing factor in these cases, but this has not been specifically reported in the literature.

There are also socioeconomic consequences that are psychosocial in nature, which are potentially just as, if not more, ruinous for the deaf individual (as interpreted by Winn et al., 2015). These include a tendency to withdraw from social interactions and also social isolation, which both impinge on quality of life and even general wellbeing (Weinstein & Ventry, 1982; Strawbridge et al., 2000; Demorest & Erdman, 1987; Grimby & Ringhdahl, 2000; Pichora-Fuller et al., 2015; Edwards, 2007). Poor compliance with the assistive hearing technology is another issue reported (such as the hearing aid-in-the-drawer phenomenon) and often accompanies the social withdrawal and even the active avoidance of listening situations (Pichora-Fuller et al., 2016). What is particularly concerning about this poor compliance is that it can occur even in cases of successful hearing aid fitting, or cochlear implantation, where the deaf individual has reported satisfactory audibility and is able to achieve high accuracy in recognition and comprehension of speech (Pichora-Fuller et al., 2016).

To add insult to injury, if the hearing impaired individual tries to cope by increasing the level of LE, this by no means offers a guarantee that the atypically increased listening demand can be overcome (Pichora-Fuller et al., 2016). In fact, the extent of resource allocation for decoding the auditory input could become such that residual capacity actually becomes insufficient for even beginning to encode and comprehend what has just been heard (Wingfield et al., 2006; Rabbitt, 1968; Suprenant, 2007). At the extreme, the brain will become unable, or even unwilling, to sustain such high levels of effort, fatiguing and/or disengaging from the task at hand (Pichora-Fuller et al., 2016). This issue is exacerbated further by the fact that the everyday listening environment is seldom ideal (Mattys et al., 2012).

1.2. Understanding listening effort

Therefore, the deaf individual becomes liable to a vicious cycle of depleting their cognitive resources to no avail (Pichora-Fuller, 2016). If this vicious cycle becomes chronic, with listening in everyday activities frequently demanding more LE than the brain is able or willing to expend, the resultant consequences becomes no longer exclusive to that of drained cognitive capacity, but also begin to include the stress response (Pichora-Fuller et al., 2016; Pichora-Fuller, 2016).

1.2.2.1. The effects of stress

The human response to stress (irrespective of whether it is external environmental demand, or internal in origin) is multifaceted, encompassing numerous aspects including the physiological, cognitive and emotional (Pichora-Fuller, 2016; Lemke & Besser, 2016; Stephens & Wand, 2012). The stress response is defined to be the point at which the individual's appraisal of the external demand appears to exceed, or strain, the individual's self-belief of ability and capacity in terms of cognitive resources (Stokols, 1992; Lazarus & Folkman, 1984; Schneiderman et al., 2005). Four main characteristics are postulated to be required within the demanding situation before a stress response is likely to be triggered: novelty, unpredictability, threat to self, and a sense of a lack of control (Lupien et al., 2012). Regrettably, these four elements can be found within the case of a challenging listening condition, especially if cognitive capacity has already been challenged by the atypical input-related demand of hearing loss (Pichora-Fuller, 2016). This then raises the alarming realisation that LE itself could become a stressor in its own right (Pichora-Fuller, 2016; Mattys et al., 2012).

This then implicates deleterious consequences for health, both physical and mental. This is because of the multidimensional aftermath of the stress response, especially when the exposure to the stressor has become repeated, or continuous (Pichora-Fuller et al., 2016; Pichora-Fuller, 2016).

In the first instance when the stressor is identified or realised, acute stress is evoked during the stage sometimes known as "alarm" (Selye, 1956; Schneiderman et al., 2005). At this point, an all-consuming cascade of changes occurs within multiple biological systems: nervous; endocrine; cardiovascular; respiratory; hepatic; renal; digestive; reproductive; and immune, all with the ultimate purpose to produce a helpful and adaptive response to the stressor (Schneiderman et al., 2005).

1.2. Understanding listening effort

To enable this valuable adaptive response, stress hormones (such as cortisol, chromogranin A and α -amylase) are initially released, one of their primary functions being to release the metabolic stores for the body's immediate use, at which point energy redistribution can occur (Kramer et al., 2016).

This energy mobilisation is diverted to the essential tissues, primarily the skeletal muscles and the brain (Schneiderman et al., 2005; McKlveen et al., 2015). Cells of the immune system (such as the macrophages) are also activated, departing from the lymphatic tissue and spleen and entering the bloodstream to migrate to tissues that are most likely to suffer damage from any kind of physical confrontation, such as the skin (Dhabar & McEwen, 1997; Segerstrom & Miller, 2004). Less critical activities are also suspended, such as digestion and the production of growth and gonadal hormones. Ultimately, these changes are meant to facilitate activities or behaviours that are not typical of daily life, in order to combat or avoid (and ultimately eradicate) the stressor (Selye, 1956; Stephens & Wand, 2012).

Two systems have been implicated in the production of the necessary stress hormones: the sympathetic nervous system (which is part of the autonomic nervous system); and also the hypothalamic-pituitary-adrenal (HPA) axis (McKlveen et al., 2015; Ulrich-Lai & Herman, 2009). The sympathetic nervous system stimulates the adrenal medulla (within the adrenal glands seated on top of the kidneys) to produce catecholamines (such as adrenaline). In parallel, the paraventricular nucleus of the hypothalamus produce corticotrophin-releasing factor, which in turn stimulates the pituitary to produce adrenocorticotrophin, which then stimulates the adrenal cortex (Schneiderman et al., 2005).

The adrenal cortex (also within the adrenal gland) subsequently secretes cortisol, a type of glucocorticoid (Stephens & Wand, 2012). Together, the action of catecholamines and cortisol have been established to promote lipolysis and gluconeogenesis, thereby yielding the required glucose for the energy demands, as well as a concert of other regulatory functions (Cohen et al., 2007).

The autonomic nervous system provides the most immediate response, with the sympathetic branch of the autonomic nervous system being engaged within seconds of stressor onset, in order to prepare the individual to respond immediately, to then rapidly subside as the result of the reflex action of the parasympathetic arm of the autonomic nervous system (McKlveen et al., 2015). An example of these rapid alterations in physiological states is increased blood pressure as well as region-specific vasodilation and vasoconstriction, helping to shunt blood to where it is needed (Ulrich-Lai & Herman, 2009).

1.2. Understanding listening effort

These initial responses can appear to be primitive, and this is because they actually are primeval in origin (Segerstrom & Miller 2004). Modern humans rarely encounter many of the stimuli that had commonly evoked the “fight-or-flight” responses within their ancestors, such as predation or inclement weather without protection. However, the human physiological response still continues to reflect the demands of earlier environments encountered within evolution, meaning that threats that do not require the primeval physical response (such as psychogenic stressors) can still induce these primitive physiological processes (Segerstrom & Miller, 2004).

This initial autonomic activation (which is extremely rapid) is followed by the activity of the HPA axis (which occurs over a slower timescale), which then regulates and sustains the glucocorticoid levels according to the energy demand (Stephens & Wand, 2012; Ulrich-Lai & Herman, 2009). It also assists in providing sophistication to the behavioural branch of the stress response, by incorporating the environmental and emotional context in which the stressor is occurring. Indeed, this is where psychogenic stressors (both real and anticipated) can begin to impart their influence (Ulrich-Lai & Herman, 2009). This is because the HPA axis receives afferent projections (at the level of hypothalamus) from a diverse group of regions such as limbic structures (including the amygdala, amygdala and prefrontal cortex); midbrain; and brainstem nuclei (Stephens & Wand, 2012; Ulrich-Lai & Herman, 2009; McKlveen et al., 2015).

These regions, in their turn, receive associational information from subcortical and cortical areas that are involved in higher-order sensory processing (such as the piriform cortex and insular cortex) and also memory (i.e. regions including the entorhinal cortex and cingulate cortices) (Ulrich-Lai & Herman, 2009). Attention and arousal can be involved in the modulation too, via connections projecting from the locus coeruleus and raphe nuclei (Ulrich-Lai & Herman, 2009). Additionally, reward processing has a role (via extensive interactions with the nucleus accumbens and ventral tegmental area), with exposure to reward buffering the effect of the stressors (Ulrich-Lai & Herman, 2009). As a result, the stress response can become tailored with respect to prior experience, as well as anticipated outcomes, with the prefrontal cortex believed to be the overall coordinator, orchestrating an integrated response (McKlveen et al., 2015).

All these interactions are mediated by a variety of brain signaling mechanisms, with the neurotransmitters involved being either inhibitory in nature (such as GABA, or opioids) or excitatory (e.g. norepinephrine and serotonin) (Stephens & Wand, 2012). Thus, the central nervous system and the endocrine system are tightly interconnected, with the possibility of tuning from the HPA axis (and also the autonomic system) so that resultant activations are in accordance with stressor modality or intensity (Ulrich-Lai & Herman, 2009).

1.2. Understanding listening effort

Overall, human beings have great resilience, being generally able to adjust to adverse situations and cope well (Schneiderman et al., 2005; Stephens & Wand, 2012). Indeed, this conglomeration of acute stress responses do not typically impose a health burden, with self-regulating allostatic processes having been evolved to maintain homeostasis and protect the organism (Selye, 1956). However, if the threat is unrelenting inducing chronic stress levels, the stress response of the body is at risk of becoming maladaptive (Schneiderman, 1983; Selye, 1956). After all, mounting a stress response is energetically demanding and can only be maintained by the body for a finite period of time (Ulrich-Lai & Herman, 2009).

Unfortunately, it has been reported by the hearing impaired populations that feelings of chronic stress can be experienced in response to listening being taxing, especially if the challenge has persisted for a sustained period of time or is frequent in occurrence (Hetú et al., 1988). It has even been postulated that social withdrawal in adulthood is actually a protective coping mechanisms adopted in response to the exposure to high and persistent stress levels (Sandi & Haller, 2015; Pichora-Fuller, 2016).

When stress has become chronic, and the body's responses have been rendered maladaptive, this is point at which the aforementioned deleterious consequences on health begin (Stephens & Wand, 2012). One of the starting points for the perpetration of this maladaptation includes the induced chronic stress triggering shifts in the normal circadian rhythm of cortisol release, as well as augmentations in cortisol levels (Bess et al., 2016; Stephens & Wand, 2012). Compellingly, neuroendocrine research conducted in deaf children has revealed elevations in the cortisol awakening response at the beginning of a school day (Bess et al., 2016).

The cortisol awakening response is a well-defined phenomenon in healthy humans and refers to the early morning rise in cortisol level upon awakening, whereby cortisol increases twofold within the first 30 to 45 minutes after waking, to then remain elevated for up to an hour (Bess et al., 2016). This response is separate to the basal diurnal fluctuations of cortisol secretion, and is often considered to be indicator of the reactivity capacity of the HPA axis (Kudielka et al., 2009). Elevations in this cortisol awakening response have already been previously associated with the individual experiencing unusual stress (Schlotz et al., 2004; Deluca, 2005; Whitehead et al. 2007; Fries et al., 2009; Kumari et al., 2009). Elevated cortisol levels have also been exhibited in some research studies investigating burnout in adults, a condition denoted by fatigue, loss of energy, and poor coping skills (Kudielka et al., 2009). Thus, elevated cortisol awakening responses are thought to be an early indicator of impending exhaustion, when the individual becomes no longer able to cope with the stressor (Bess et al., 2016).

So, for these cortisol elevations to also be present in deaf children, there is the implication that their bodies and brains may be sensing (consciously and/or unconsciously) an impending threat in the upcoming school day and mobilising energy stores (as well as enhancing vigilance levels) accordingly in preparation (Bess et al., 2016). These elevations in cortisol level are a particular concern because it is now known that too much exposure to cortisol is detrimental to both the body and brain (Stephens & Wand, 2012; De Vente et al., 2003; Grossi et al., 2005; Kudielka et al., 2006).

There is a negative feedback loop, which is carefully designed to control the levels of hormones (such as that of cortisol) in order to maintain the critical homeostasis (Stephens & Wand, 2012). This mechanism usually ensures that the levels of cortisol (and other hormones) are maintained at their required levels to enable optimal bodily function (Stephens & Wand, 2012; Bess et al., 2016). This compensatory process becomes especially pivotal when the body and/or brain is placed under duress, such as that of the presence of a stressor (Stephens & Wand, 2012).

However, this negative feedback loop is incredibly fragile and thus relatively prone to disruption, such as that of excessively raised cortisol levels (Stephens & Wand, 2012). Thus, the interference of the fine balance in the interactions of cortisol, with both the HPA axis and the sympathetic nervous system, leads to cumulative increases in allostatic costs that end up leading to severe (and even intractable) physical and psychological ramifications. (McEwen, 1998; Schneiderman et al., 2005). This allostatic toll comprises a myriad of damaging physical and psychological symptoms (Stephens & Wand, 2012).

For instance, within the physical domain, chronic sympathetic nervous stimulation of the cardiovascular system can lead to hypertension and even vascular hypertrophy, which could then eventually lead to damage of the ventricles and arteries (Henry et al., 1975; Schneiderman et al., 2005). This then imparts increased risk of hypertension; cardiovascular disease; heart attacks; and even strokes and respiratory failure (Brownley et al., 2000; Spruill, 2010).

Additionally, both humoral and cellular immunity can become suppressed, via mechanisms such as the elevated stress hormones affecting cytokines (which are the primary communication molecules produced by immune cells) which then has ramifications for wound healing and the ability to fight off infections (Segerstrom & Miller, 2004; Kiecolt-Glaser et al., 2002).

1.2. Understanding listening effort

Not only this, the stressor continues to promote proinflammatory cytokines indefinitely, thus inducing vulnerability to autoimmune disease and myalgic encephalomyelitis (or chronic fatigue syndrome), as well as joint pain; coeliac disease: and irritable bowel syndrome (Jerjes et al., 2005; Roberts et al., 2009; Nijhof et al., 2014; Parker et al., 2001; Bess et al., 2016; Hardy & Tye-Din, 2016; Videlock et al., 2016). Infertility can also be included within the physical symptomatology (in both males and females) due to disrupted function, or even complete suspension, of gonadal hormones (Schneiderman et al., 2005; Kato et al., 2013).

The psychological symptoms are equally diverse in their manifestation. For instance, the unruly inflammatory processes can adversely affect mental wellbeing, with malaise; diminished appetite; dysphoria; and also depression being induced (Schneiderman et al., 2005). Not only this, chronic stress has also been associated with anatomical changes within the brain at multiple levels: synaptic, neuronal, and network (Schneiderman et al., 2005; Ulrich-Lai & Herman, 2009). For instance, within the limbic structures of the hippocampus and prefrontal cortex, there have been reports of reduction in dendritic branching and spine density, which would then confer altered neuronal communication capacity (Magarino & McEwen, 1995; Radley et al., 2008; Ulrich-Lai & Herman, 2009).

Altered expression of neurotransmitter receptor subunits has been observed too (as well as reports of abnormal levels of the GABA neurotransmitter within the HPA axis itself), which would then further disrupt the complex cascade of neuronal communication (Cullinan, 2000; Ziegler et al., 2005; Bowers et al., 1998; Ulrich-Lai & Herman, 2009).

Indeed, at the functional level within the prefrontal cortex, the interference with the neuronal activity is such that the prefrontal cortex appears to have been rendered “offline” (Stephens & Wand, 2012; Ulrich-Lai & Herman, 2009; Diamond et al., 2007; McKlveen et al., 2015). Because of the recognised criticality of the prefrontal cortex in higher-order cognitive functioning (with it being posited as the coordinator of stress responding, as well as being key to cognitive control), this then immediately implicates compromised cognition (Ulrich-Lai & Herman, 2009; Diamond et al., 2007; Liston et al., 2006; McKlveen et al., 2015).

Consistent with this, a proliferation of cognitive disturbances has been reported in stress research literature (Schneiderman et al., 2005; Stephens & Wand, 2012; Ulrich-Lai & Herman, 2009).

1.2. Understanding listening effort

For instance, inappropriate processing of stress information is an issue, with cortisol stipulated to dysregulate interactions with regions such as the nucleus accumbens and ventral tegmental area and their interactions with limbic structures such as the prefrontal cortex, thus compromising reward processing (Marinelli & Piazza, 2002; Casey et al., 2008; Stephens & Wand, 2012).

Emotional reactivity becomes volatile too, with the amygdala (another limbic structure) affected by cortisol's disruption and affective processing subsequently interfered with (Stephens & Wand, 2012; Casey et al., 2008). The dysregulated reward and affective processing immediately have implications for any kind of cognitive appraisal regarding self-ability and capacity in meeting any kind of input-related demand (Ulrich-Lai & Herman, 2009; Stephens & Wand, 2012).

In addition, the malfunctioning reward and affective systems lead to increased likelihood of unhealthy reward-seeking behaviour (such as intake of palatable food, or seeking pharmacologically-induced euphoria) (Ulrich-Lai & Herman, 2009). There then accrues an increased risk of obesity and substance abuse (such as alcoholism) (Stephens & Wand, 2012).

Just as optimal levels of cortisol are required for healthy physical function, optimal levels of cortisol are also posited to be needed for learning and memory (Stephens & Wand, 2012). Indeed, cortisol has been implicated in the phenomenon of one of the potentially key mechanisms for learning and memory: long-term potentiation (which is involved in enhancing synaptic transmission) (Goosens & Maren, 2002; Stephens & Wand, 2012). This is because a wide distribution of glucocorticoid receptors (including those for cortisol) has been discovered not just in the hypothalamus, but also above it in regions such as the hippocampus and prefrontal cortex (Stephens & Wand, 2012). These are affected by altered cortisol levels and, consequently, profound memory impairments have been associated with chronic stress (Pichora-Fuller, 2016; Diamond et al., 2007).

In addition, aberrant biases in emotional learning and habit-based forms of learning (mediated by regions including the amygdala and dorsal striatum) are also often observed, and frequently occur in lieu of essential cognitive control (i.e. goal-directed performance and appropriate allocation policy of cognitive resources) (Stephens & Wand, 2012).

This extensive disruption of cognitive functioning will inevitably impact on the ability of the deaf individual to meet any input-related demand provided by a given listening situation (Pichora-Fuller, 2016; Pichora-Fuller et al., 2015).

This, in turn, may continue to elevate LE as the individual tries to achieve the desired performance level (Pichora-Fuller et al., 2016).

Unfortunately, there is also the potential for mental health deterioration (Stephens & Wand, 2012). The neurological changes generated by chronic stress can become sufficiently pervasive and destructive that psychiatric disorders can develop (Kato et al., 2013). For example, it is extensively reported that the first clinical episode of depression often develops following a major negative life event (Paykel, 2001; Hammen, 2005; Schneiderman et al., 2005; Stephens & Wand, 2012). Clinical anxiety is another reported consequence of a stressful life event (Faravelli & Pallanti, 1989; Finlay-Jones & Brown, 1981).

This potential for mental health breakdown is concerning, particularly because it is already well established that there is an increased prevalence of psychiatric disorders in the deaf population (Du Feu & Fergusson, 2003). In the case of deaf children, there is an increase of 20% in risk relative to normal hearing children for mental health problems (Hindley et al., 1994; Hindley, 2000; Du Feu & Fergusson, 2003).

This susceptibility to mental health disorders is not unique to the cases of early onset or congenital deafness, with cases of intermittent deafness reporting issues such as symptoms of paranoia and also depression (Eastwood et al., 1985; Du Feu & Fergusson, 2003). The cumulative effects of a progressive hearing loss renders the individual vulnerable to mental health problems too (Du Feu & Fergusson, 2003). Another corollary that has been investigated is the possibility that the increased cognitive load of listening could even be a factor mediating the relationship between hearing loss and cognitive decline and/or dementia (Lin & Albert, 2014; Uhlmann et al., 1989; Lin et al., 2011, 2013; Gallacher et al., 2012).

A commonly stated putative mechanism for the pathophysiology of psychiatric disorders (such as major depression, clinical anxiety, post-traumatic stress disorder, schizophrenia, and even borderline personality disorder, paranoia and psychosis) is the dysregulation of the HPA axis (and its interactions with the central nervous system thereupon) (Roy et al., 1988; Lesch et al., 1990; Holsboer, 2000; Kunugi et al., 2006; Stephens & Wand, 2012; Schneiderman et al., 2005; Carvalho Fernando et al., 2012; Kato et al., 2013; Burmeister et al., 2008).

Mechanisms such as aberrant cortisol levels leading to disrupted neurotransmitter signalling and malfunctioning neuronal networks are also frequently postulated to underlie the imbalanced HPA axis activity (Schneiderman et al., 2005; Stephens & Wand, 2012; Roy et al., 1988; Lesch et al., 1990). A concomitant pathway stipulated is also that of microglial action (Kato et al., 2013).

The involvement of microglial action has been investigated because it is believed to bridge the endocrine and the nervous systems together, enabling interactions (Kato et al., 2013).

Microglia are often conceptualised as the macrophages of the central nervous system, i.e. becoming the kingpins of neuro-immunology (Del Rio-Hortega, 1919; Kato et al., 2013). In particular, microglia are thought to be particularly key for the neuro-inflammatory pathways underlying the neuro-immunological system, being responsible for monitoring micro-environmental changes and releasing cytokines to promote inflammatory reactions as required (Block et al., 2007; Kato et al., 2013).

Not only this, cortisol and also corticotrophin-releasing factor (secreted by the HPA axis) have already been shown to dysregulate microglial action, such that any kind of central nervous inflammation (that is a natural part of an immune response) propagates uncontrollably leading to neurodegeneration (Kato et al., 2013).

Consistent with the proposed significance of this microglial machinery (in the pathophysiology of psychiatric disorders) are the findings of recent positron emission tomography studies, which revealed abnormal microglial activity in disorders such as schizophrenia, depression and even autism (Steiner et al., 2006, 2008; van Berckle et al., 2008; Doorduyn et al., 2009; Morgan et al., 2010, 2012; Kato et al., 2013).

To ramify any kind of physiological and psychological instability in times of chronic stress, there is also the complex interplay between genes and the environment complicating matters (Stephens & Wand, 2012). An example of this is epigenetics (Kato et al., 2013; Lee et al., 2010; Stephens & Wands, 2012). Epigenetics is a wide-ranging term encompassing a number of different mechanisms (such as DNA methylation and histone modification) which are essential to genomic stability and gene expression, particularly in the guidance of the activation and silencing of specific genes at certain times in response to environmental factors (Provenzano & Domann, 2007).

An overwhelming amount of literature has labelled epigenetics as the perpetrator in a wide range of pathology, with the potency of inappropriate or misdirected epigenetic action such that it could even induce the pathogenesis of cancer (Provenzano & Domann, 2007). It has been proposed that any inherent genetic flaws, or weaknesses, which naturally occur within the human can become amplified in the case of chronic stress, because epigenetic mechanisms could inadvertently render inappropriate mutational events in response (Provenzano & Domann, 2007; Stephens & Wand, 2012; Kato et al., 2013; Niwa et al., 2013). The individual differences in genetic code vulnerable to this unintentional amplification include that of HPA axis reactivity (Stephens & Wand, 2012; Bess et al., 2016).

Not only this, there are countless studies suggesting that genetic vulnerabilities can also predicate susceptibility to psychiatric disorders (Burmeister et al., 2008; Casey et al., 2008).

Indeed, the hereditary basis for psychiatric disorders has been known from the beginning of the nineteenth century (Zec, 1995). There has been the interesting notion that these genetic vulnerabilities to psychiatric disorders (such as that of anxiety, bipolar disorder, depression, schizophrenia, panic disorder, and psychosis) actually emerge at the extreme ends of the normal population variation of personality (e.g. neuroticism) and volition (Burmeister et al., 2008; Kato et al., 2013).

This, therefore, has the distressing consequence that a full-blown psychiatric disorder could be generated from an inherent susceptibility, or even personality trait (Provenzano & Domann, 2007; Niwa et al., 2013; Kato et al., 2013). In line with this notion, research using genetic animal models has indicated that the combination of stress and cortisol (and other glucocorticoids) are capable of altering the methylation patterns of other genes to then invoke psychotic depression and schizophrenia (Niwa et al., 2013; Kato et al., 2013; Stephens & Wand, 2012).

To even further exacerbate matters, there is also perturbing neurochemical evidence to suggest that chronic stress could also cause the sensitisation of the stress response, meaning that there are excessive responses of the autonomic-neuroendocrine system to external stimuli and also the internal cognitive appraisal of the stimuli (Ulrich-Lai & Herman, 2009). Not only this, the allostatic injury inflicted by these sensitised stress responses becomes even greater than before (McEwen & Gianaros, 2010; Stephens & Wand, 2012).

Thus, an inexorable cycle of even more harm can be induced within the physiological and psychological milieu due to the recoil of the human body and the brain to the burden of chronic stress.

In summary, bearing in mind that it is conceivable that a stress response can be triggered by listening becoming too effortful, a constellation of exaggerated and essentially inappropriate physiological reactions can arise from what is simply a challenging listening situation (Pichora-Fuller, 2016). The resultant allostatic injury is to the profound detriment of the individual, compromising both physical and mental health and wellbeing (Stephens & Wand, 2012; Schneiderman et al., 2005).

Not only this, in addition to a drained cognitive capacity triggering the stress response, stress itself can also weaken cognitive function needed for comprehension, as well as potentially distort cognitive appraisal (Pichora-Fuller, 2016; Lemke & Besser, 2016).

1.2. Understanding listening effort

This then increases the possibility of maladaptive strategies being implemented in allocation policy, as well as enhance the likelihood of triggering the stress response yet again, with its associated cascade of devastation (Pichora-Fuller, 2016).

All of this ultimately culminates in the further fuelling of the aforesaid vicious cycle that the deaf individual may have entered in, i.e. the unproductive draining of the finite pool of cognitive resources to no avail in listening performance (Pichora-Fuller, 2016; Lemke & Besser, 2016).

It is already alarming that such penalties to general health and wellbeing can be accrued (via a completely involuntary and primitive cascade of physiological events), and this is even before considering the additional physiological and psychological consequences that are unique to the event of a brain becoming deafened (McGarrigle et al., 2014; Kral et al., 2016).

1.2.2.2. The effects of the connectome disease

Hearing impairment is not merely just compromised of sensory transduction (McGarrigle et al., 2014; Kral et al., 2016). A miscellany of deleterious physiological and psychological consequences has been implicated in the event of the brain becoming deafened (Lazard et al., 2012b; Kral et al., 2016).

In fact, deafness has recently been christened as a “connectome disease” (O’Donoghue et al., 2016; Kral et al., 2016). The connectome is the network of effective synaptic connections and neural projections that shape global communication and form the self-organising brain capable of dynamic interaction with the environment (Kral et al., 2016; NIH Blueprint, 2010; Sporns, 2015; Hübener & Bonhoeffer, 2014), with each individual’s wiring within their own connectome being unique to the individual (O’Donoghue et al., 2016). When considering just the auditory component of the central nervous system alone, a phenomenal assembly of connections arise. This is because the auditory component of the human connectome serves as a fundamental principle to guide the organisation and sequencing of behaviour and cognition, especially for neurocognition (Kral et al., 2016).

For instance, within the memory domain, the auditory system’s multitude of interactions is thought to subsume the substrate for implicit memory (which involves brain regions such as the basal ganglia and cerebellum), as well as contribute to explicit declarative memory and spatial orientation, thus involving the entorhinal cortex and hippocampus (Weinberger, 2011; Edeline, 2012). Fear memory is also of importance, which is governed by regions such as the amygdala (Kral et al., 2016).

1.2. Understanding listening effort

In addition, in order to embrace the facet of language processing, these memory subcomponents are also believed to include semantic memory (enabling access to stored linguistic knowledge, i.e. vocabulary, as well as phonological and syntactic representations) and also episodic memory (required to relate incoming information to past episodes of personal experience, which includes specific conversations and specific social interactions) (Lemke & Besser, 2016). The crucial process of working memory must also be included as well (Baddeley, 1992; Lemke & Bessser, 2016). Working memory is generally defined as the temporary storage for information that enables online encoding and decoding, as well as the relation of incoming auditory information to representations of facts and episodes and linguistic components stored in semantic and episodic memory (Daneman & Merikle, 1996; Rönnberg et al., 2013).

Working memory, in turn, by being a subclass of this family, implicates executive function, an umbrella term that itself comprises an even more extensive neural network believed to enable cognitive control and oversight processes required to undertake planned and goal-directed activities (Barkley, 2012; Diamond, 2013).

The many components of executive function cited in cognitive psychological and behavioural research have included attention, inhibitory and interference control, sequential processing, concept formation, processing speed, and also cognitive flexibility (Kahneman, 2011; Dye, 2014; Buckley et al., 2010; Kronenberger et al., 2013; Lyxell et al., 2008).

These components ultimately lend to the individual being able to remain focused and quickly (and flexibly) adapt to the constantly changing environment, which is particularly relevant during the fast discourse that is characteristic of spoken communication, which requires ongoing planning, monitoring, evaluating, and reasoning to guide thought, response and intention (Miller & Cohen, 2001).

Therefore, there is a decidedly dynamic, reciprocal and multimodal interaction of language, sensory experience and neurocognition to produce the holistic functioning brain, with each aspect providing scaffolding for each other (Figueras et al., 2008; Conway et al., 2009; Lazard et al., 2012b).

Thus, when sensory deprivation strikes, there is collateral damage (beyond that of just compromised auditory perception) on a global scale (O'Donoghue et al., 2016; Kral et al., 2016). To name just a few, phonological processing, verbal working memory and cognitive fluency are at risk in deaf individuals (Rudner et al., 2010; Nitttrouer et al., 2013; Rönnberg et al., 2008; Barkley, 2012; Kronenberger et al., 2013, 2014; Wass et al., 2008; Lyxell et al., 2008).

1.2. Understanding listening effort

This ripple effect of disruption can only be aggravated if sensory loss has occurred at a time when the brain is in a highly vulnerable state, such as the juvenile brain during both early neurodevelopment and also adolescence. This is because juvenile neurons are more susceptible to modification in their mutual interconnections (in response to changes in environmental stimulation), and this includes undesired changes that deviate from the original blueprint determined by the human genome (Kral et al., 2016; Hübener & Bonhoeffer, 2014; Whiteus et al., 2014).

These aberrations include delayed cortical synaptogenesis; indiscriminate synaptic pruning; as well as functional immaturity (Kral & Sharma, 2012; Kral, 2013; Kral et al., 2005, 2016). Since the purpose of these mechanisms is to promote appropriate development of vital central auditory processes and cortico-cortical interactions, already present disruptions in auditory processing (such as that needed for auditory object perception) would unduly ramify (Kral & O'Donoghue, 2010; Kral et al., 2016; O'Donoghue et al., 2016).

With the human cochlea operational by 24 weeks post conception, the toll of deafness can manifest even in utero (Kral et al., 2016). In fact, deafness via loss of cochlear cells can still intrude on subsequent neurodevelopment due to increased likelihood of apoptosis of auditory brainstem neurons (Tong et al., 2015; McBride et al., 2013).

All of these physiological events are essentially deleterious due to the postulated role of auditory experience in providing temporal patterns to the brain, which is then essential for sequential processing abilities within not only the sensory domain, but also in neurocognition, such as pattern detection, sequential memory, and even sustained attention (Conway et al., 2009, 2011a, b).

Despite such abnormal maturation, a level of residual plasticity does still remain (Kral et al., 2002; Schramm et al., 2002; Sharma et al., 2007, 2009). In fact, the deafened brain has been stated to be capable of developing enhanced (or even superior) faculties, such as that of spatial attentional capacities (Dye & Bavelier, 2010). These enriched functions are believed to partially arise from the potentiation of cortical areas that are already multimodal in design, due to its inputs being more multisensory in nature (Fine et al., 2005; Bavelier et al., 2006).

This residual facilitative plasticity, however, is thought to be undermined by functional decoupling, whereby the primary auditory cortex is no longer capable of being modulated by higher auditory fields (Kral & Sharma, 2012).

At the extreme, there may even be maladaptive plasticity via the colonisation of other sensory modalities (e.g. vision) within the brain regions designated for auditory processing, thus introducing perceptual integration conflicts when auditory stimulation is provided, especially if deafness has occurred within the genetically pre-determined auditory and language sensitive periods (Champoux et al., 2009; Finney et al., 2001; Auer et al., 2007; Lazard et al., 2012a, b; Lyness et al., 2013; Chun et al., 2013; Blundon & Zakharenko, 2013).

Thus, the nature of the adaptive plasticity in response to hearing impairment (facilitative or maladaptive), and the extent to which it pervades the neuronal networks, ultimately affect the brain's ability to encode and process sound (O'Donoghue et al., 2016; Kral et al., 2016).

Unfortunately, the adolescent is no less vulnerable to the devastation of deafness. Despite a considerable degree of maturation having been achieved by adolescence, research has recently revealed a surprising fragility and lability to the adolescent brain (Steinberg, 2005; Casey et al., 2008).

Increased emotional reactivity and exaggerated cognitive appraisal of both positive and negative rewards are just two of the reported characteristics of adolescent neurocognition (Pine et al., 2001; Silveri et al., 2004; Steinberg, 2005; Casey et al., 2008). It is thought that, because the prefrontal cortex is one of the last brain regions to fully mature, that there is still an insufficiency in cognitive control of the relatively more mature limbic system, especially with regards to reward and affective processing in the presence of compelling incentives and also social peer pressure (Casey et al., 2000, 2002, 2005a, b; Blakemore, 2008; Galvan et al., 2006; Monk et al., 2003; Thomas et al., 2004). This notion has been supported by evidence provided by structural magnetic resonance imaging and diffusion tensor imaging, with this altered connectivity, as well as changing ratios of grey and white matter, being observed (Casey et al., 2008).

To compound matters, this imbalance in maturational trajectories of the different brain regions coincides with puberty, during which significant and volatile neuro-endocrinological changes occur, such as the increases of adrenal and gonadal hormones, both of which have pervasive influences on brain function (Spear, 2000; Casey et al., 2008; Blakemore, 2008; Arnsten & Shansky, 2004; Sisk & Foster, 2004; Steinberg, 2005). Thus, any sensory (and associated cognitive disruption) unduly ramify.

1.2. Understanding listening effort

Therefore, it can be concluded that deafness is not merely the disease or dysfunction of the ear, it constitutes a cascade of disturbance of the entire connectome beyond that of hearing, especially if auditory deprivation occurs at a vulnerable time in the brain's maturation.

This pervasive disruption can only further compound the demand for cognitive capacity and render the vicious cycle of cognitive depletion and also the allostatic toll of chronic stress all the more injurious (Kral et al., 2016; O'Donoghue et al., 2016; Pichora-Fuller, 2016; Pichora-Fuller et al., 2016).

It is essential to recognise that all these connotations of deafness as a connectome disease will also apply to the CI user population, however successfully they appear to be using their implant (Kral et al., 2016; O'Donoghue et al., 2016).

To add another layer of complexity for CI users, there is also the reality of electric hearing, the consequences of which must not be overlooked (Kral et al., 2016; O'Donoghue et al., 2016).

1.2.2.3. The reality of electric hearing

The cell bodies of the spiral ganglion in the inner ear are remarkably robust, being able to withstand prolonged deafness, as well as virulent etiologies such as meningitis (Hinojosa & Marion, 1983; Miura et al., 2002; Leake & Rebscher, 2004). Thus, even in cases of total deafness, direct electrical stimulation (via the nodes of Ranvier of these cell bodies) can be used to evoke responses within the auditory nerve (Wilson & Dorman, 2008a, b). This electrical stimulation is delivered by the CI neuroprosthesis via an array of 12-22 electrodes that have been surgically inserted in the scala tympani (i.e. one of the three fluid-filled chambers of the cochlea: Wilson & Dorman, 2008a, b; Finley & Skinner, 2008; Escudé et al., 2006; Lazard et al., 2012a, b). A spatiotemporal code can then be elicited to some degree within these auditory nerve fibre responses, thus generating auditory sensation (Tillein et al., 2015; Hartmann & Kral, 2004; Wilson & Dorman, 2008a, b; Kral et al., 2016).

This mode of electrical stimulation differs from the simple amplification of acoustic signals that is performed by conventional hearing aids, and is remarkably effective (Faulkner & Pisoni, 2013; Lazard et al., 2012a, b; Kral et al., 2016; Peterson et al., 2010). However, despite much progress in device design, processing strategies, noise reduction algorithms and surgical techniques, there are certain limitations in what the CI is able to achieve (Wilson, 2015; Wilson & Dorman, 2008a, b).

1.2. Understanding listening effort

For instance, there has been an ongoing debate about the number of channels (i.e. provided by the electrode array) that can actually be utilised in sound processing. A series of studies have suggested that there are no more than 4-8 functionally independent channels (Lawson et al., 1996; Fishman et al., 1997; Wilson, 1997; Kiefer et al., 2000; Friesen et al., 2001; Garnham et al., 2002). However, data from a recent study suggest otherwise, arguing that the full range of 22 electrodes actively contributes to improving speech understanding (Croghan et al., 2017). Nevertheless, even if the cochlear implant's full functional capacity is available, the spatial specificity (of the auditory input) typically achieved by the CI device is markedly lower than those demonstrated in neural tuning curves obtained in the normal hearing (Van den Honert & Stypulkowski, 1987).

These putative mechanisms underlying the CI's limitations are believed to arise from multiple domains. However, there is still ongoing research to fully elucidate the extent and nature of their impact. Such domains include the physiology of human hearing and auditory perception, as well as technical factors encompassing the CI technology itself and also surgical technique (Wilson & Dorman, 2008a, b; Faulkner & Pisoni, 2013).

With regards to the CI technology itself, there are key elements in how it processes sound that are fundamentally different to the actual physiology of human hearing (Wilson, 2015). The CI collects sound (via the microphone of the external processor) and converts it into electric signals, which are then processed through a series of bandpass filters according to the specific parameters of the CI's processing strategy (Peterson et al., 2010; ASHA, 2004). These filtered signals are then transmitted transcutaneously (i.e. across intact skin) to the implanted device, where they are converted into a series of envelope electrical signals that are carried down to the electrode array in accordance to the original bandpass filtering (Peterson et al., 2010; ASHA, 2004).

Via the longitudinal placement of electrodes along the CI's array, the CI device attempts to exploit the intrinsic tonotopy of the cochlea (i.e. basal regions of the cochlea represent higher frequency sounds, whilst more apical regions of the cochlea encode lower frequency sounds) (Kral et al., 2016). Whilst different electrodes of this implanted array may stimulate different subpopulations of neurons in a tonotopic manner, the CI's processing still omits the detail of the travelling wave of mechanical displacements within the basilar membrane in response to acoustic stimuli (Robles & Ruggero, 2001), as well as the spatial sharpening of this membrane response, via the active processes at the outer hair cells (Robles & Ruggero, 2001; Dallos, 1992).

There is also the absence of the different types of compression of neuronal signals at the synapses, such as those between the inner hair cells and the single fibres of the auditory fibres (Smith, 1985; Guinan, 1996; Kiang et al., 1988).

The influence of spontaneous nerve activity, and the broad distribution of threshold for multiple afferent fibres innervating each inner hair cell, are also not accounted for (Kiang et al., 1965; Liberman, 1978; Hartmann et al., 1984; Kiang et al., 1970; Shepherd & Javel, 1997). To further exacerbate matters, there are substantial overlaps in the electric fields generated by adjacent (and even distal) electrodes. This is because the electrodes are situated in the highly conductive fluid of the perilymph in the scala tympani (Wilson & Dorman, 2008a, b).

There has been an active attempt to minimise these overlapping electric fields, such as the development of processing strategies deploying specific patterns in the nature and timing of the electrical pulses delivered by the electrodes, as well as the deliberate deactivation of problematic electrodes (Wilson & Dorman, 2008a, b). However, this does not change the fact that the CI has poorer spectral resolution and also a reduced dynamic range compared to normal hearing (Gifford et al., 2011; Garadat et al., 2013; Srinivasan et al., 2013; Bierer et al., 2010; Faulkner & Pisoni, 2013; Hartmann & Kral, 2004).

These CI properties also hinder more complex auditory processing, such as music perception or speech understanding in adverse listening conditions (e.g. presence of noise, or telephone communication) (Shannon et al., 1995; Friesen et al., 2001, Smith et al., 2002; Faulkner & Pisoni, 2013; Caldwell & Nittrouer, 2013; Stickney et al., 2006; Schafer & Thibodeau, 2004).

Not only are there limitations within the device, there also appear to be limitations in the ability of the brain to actually utilise the incoming information produced by the electric hearing. For example, in the case of temporal fine structure (a cue deemed to be critical for speech reception, and also for music processing and sound lateralisation), the phenomenon of a “pitch saturation limit” has been reported in CI users (Smith et al., 2002; Wilson et al., 2005, 2011; Hochmair et al. 2006; Arnoldner et al., 2007; Nie et al., 2005; Zeng, 2002; Wilson & Dorman, 2008a, b). This limit is related to how changes in the rate of stimulation of a given electrode are perceived as changes in pitch (Zeng, 2002). The pitch saturation limit is observed to be around 300 Hz for most CI users, with rare cases of patients managing to achieve 1,000Hz (Zeng, 2002).

Thus, the representation of temporal fine structure information via a temporal code (such as the timing and/or frequency of pulse presentation of electrical current within the CI electrode array) cannot be fully exploited, being limited to 300Hz or lower, and may even become degraded in nature (Zeng, 2002; Baumann & Nobbe, 2004). This issue can only be aggravated by the presence of relatively sparse functional channels.

Attempts have been made to circumvent this limitation in temporal fine structure representation via the use of virtual channels (sometimes called current steering: Firszt et al., 2007; Koch et al., 2007). However, it is still not fully clear, nor completely understood, how much temporal fine structure information is actually being conveyed to the patient, irrespective of the approach taken for processing strategy (Wilson & Dorman, 2008a, b).

Unfortunately, the compromised resolution in auditory information is not necessarily improved via bilateral implantation (i.e. both ears receiving a CI). Hearing with two ears is known to provide significant benefits with regards to sound localisation, stream segregation and dereverberation (of competing sound signals) (Litovsky et al., 2012; Faulkner & Pisoni, 2013). Indeed, bilateral CI users do exhibit improvements in these aspects of auditory processing (Aronoff et al., 2010; Chan et al., 2008; Litovsky et al., 2009). Unfortunately, however, bilateral implantation does not necessarily guarantee the CI user full utility of binaural hearing cues, such as that of interaural timing differences and interaural level differences (Faulkner & Pisoni, 2013).

One possible reason for this is that the clinical and mapping strategies implemented may not be able to coordinate the inputs from the two separate CIs (Faulkner & Pisoni, 2013). In addition, there may be significant differences in depth of electrode insertion between the two ears. Not only this, the corresponding channels may also be imbalanced in terms of loudness (Goupell et al., 2013; Gordon et al., 2012; Faulkner & Pisoni, 2013). Then, to further complicate matters, there may be deafness-related disturbances, and even reorganisation, of the binaural circuitry (Smith & Delgutte, 2008). Thus, even if the two implants are well balanced and matched, the neural machinery may not be capable of utilising the available binaural time and intensity cues (Smith & Delgutte, 2008; Faulkner & Pisoni, 2013). Therefore, bilateral hearing (via two CIs) is not always synonymous with binaural hearing (Faulkner & Pisoni, 2013).

There are also surgical elements to consider. The orientation and geometric arrangement of the CI electrodes and their proximity to the target neural structures is important for optimising the impact and utility of the electrical stimulation (Wilson & Dorman, 2008a, b; Aschendorff et al., 2007; Holden et al., 2013; Skinner et al., 2007; Anderson et al., 2017a, b; Lazard et al., 2012a, b).

However, the depth that can actually be achieved in the insertion of the electrode array within the scala tympani is restricted by decreasing lumen size (from base to apex: Wilson & Dorman, 2008a, b).

To additionally complicate matters, there is also the unevenness of lumen (particularly in the apical region) and the curvature of the cochlear spiral to contend with (Wilson & Dorman, 2008a, b; Blamey et al., 1992; Finley & Skinner, 2008; Skinner et al., 2002; Yukawa et al., 2003; Anderson et al., 2017a, b; Holden et al., 2013). No array has been inserted deeper than 30mm (the total length of the cochlea is 35mm), with insertion depths typically at 18-26mm instead (Wilson & Dorman, 2008a, b). In some cases, insertions can be even shallower, due to the presence of bony obstructions (i.e. ossification) in the lumen (Wilson & Dorman, 2008a, b).

“Soft surgery” introduction techniques and the latest design of electrode arrays have endeavoured to decrease the distance between the electrode arrays and the spiral ganglion cells (Lazard et al., 2012a, b), but there are still other elements of the CI surgery that could provide obstacles to the CI’s functionality. For instance, tissue damage can result from implantation trauma, which could then trigger inflammatory responses resulting in loss of neurons (Kral et al., 2016). If cochlear health and/or nerve survival is already an issue, then the effect of any additional neuronal loss is likely to be detrimental.

The factor of nerve survival is implicated because, in the absence of the normal quantity and quality stimulation provided by the inner hair cells in the deafened cochlea, the dendrites (i.e. the peripheral components of the neurons between the cell bodies in the spiral ganglion and the terminals within the organ of the Corti) often undergo retrograde degeneration and eventually cease to function (Hinojosa & Marion, 1983). This retrograde degeneration is rarely uniform both within (i.e. certain parts of the cochlea may have more intact cells and connections than other parts) and also across patients (Wilson & Dorman, 2008a, b).

Not only this, the retrograde degeneration sometimes extends to the spiral ganglion cells themselves, as well as the axons connecting these cells to the central nervous system (Wilson & Dorman, 2008a, b; Leake & Rebscher, 2004). This has the consequence that, from patient to patient, the neural substrate available (i.e. the target for the CI) can vary considerably, ranging from sparse to substantial (Wilson & Dorman, 2008a, b; Leake & Rebscher, 2004).

Interestingly, however, the extent of nerve survival and how this relates to CI outcomes is still contentious (Wilson & Dorman, 2008a, b). Whilst it is intuitive to assume that reduced neural infrastructure to receive and convey the CI's electrical stimulation would impede functional outcomes, there is compelling evidence to suggest that this is not the case (Wilson & Dorman, 2008a, b).

Indeed, negative correlations, and even the complete absence of a relationship, have been found between the number of surviving ganglion cells and prior word recognition scores within post-mortem histological studies on donated temporal bones (containing the cochlea: Nadol et al., 2001; Khan et al., 2005; Fayad & Linthicum, 2006).

There have been striking cases where survival of spiral ganglion cells was negligible yet the deceased CI user had achieved high monosyllabic word recognition rates in life. Conversely, excellent ganglion cell survival did not guarantee high monosyllabic word recognition (Nadol et al., 2001; Khan et al., 2005; Fayad & Linthicum, 2006). Thus, whilst a degree of spiral ganglion cells is mandatory to enable function of the CI neuroprosthesis, the actual required number may only be small (Wilson & Dorman, 2008a, b).

Yet another complication regarding the insertion of the electrode array is tonotopic mismatch (Svirsky et al., 2004). This is where the placement of a specific electrode (or electrodes) responsible for a given frequency does not align with the neural subpopulations typically responsible for conveying that particular frequency (Baskent & Shannon, 2005; Svirsky et al., 2004). This would then serve to create yet another limiting factor in the frequency and intensity information available about the sound (Faulkner & Pisoni, 2013).

Overall, there is a stark contrast between the operating mechanisms, and also capability, of the CI neuroprosthesis and the complex and intricate interplay of the natural cochlear system. This is particularly bearing in mind the minute mechanical machinery of the cochlea, as well as its relationship to over 15,000 sensory hair cells that are themselves connected with 30,000 neurons (Wilson, 2015).

Indeed, the CI was considered to be such crude mimicry of this system that, when the technology was first introduced, there was considerable scepticism as to whether the neuroprosthesis would even succeed (Wilson, 2015). However, CI users have repeatedly demonstrated the ability to acclimatise to the artificial and highly degraded acoustic input and learn to use it in language development as well as everyday listening (Shannon et al., 1995; Dorman et al., 2008; Davidson et al., 2011; Kral et al., 2016; Niparko et al., 2010).

Therefore, it is testament to the brain that it is able to utilise this relatively unsophisticated and unnatural imitation of peripheral auditory physiology (Wilson, 2015).

Nonetheless, ultimately, CIs are not a cure for deafness (Faulkner & Pisoni, 2013). Thus, the aforementioned consequences of the deafened brain almost certainly apply (Kral et al., 2016; O'Donoghue et al., 2016; Lazard et al., 2012a, b).

So, the fact that CI users are capable of such excellence in listening ability and speech production (despite such sparse auditory representation) does beg the question: "At what cost?" (in terms of the resulting demand on cognitive resources).

To further exacerbate matters, it is now being recognised by the audiological profession that current clinical assessment (utilising audiometry and speech intelligibility tests) may not be sufficient to fully capture the reality of CI user performance (McGarrigle et al., 2014; Pichora-Fuller et al., 2016). In fact, a dissociation has been reported between LE and speech intelligibility (Humes, 1999; Winn et al., 2015; Pals et al., 2013; Sarampalis et al., 2009).

For instance, when principle component analyses were used to analyse hearing aid outcome to canvas its multidimensional nature, the factor of LE emerged as a separate factor to that of speech intelligibility (Humes, 1999).

This independence of LE to speech intelligibility is also yielded by studies incorporating CI simulations, at both a behavioural and a physiological level: as speech intelligibility performance plateaus beyond a given level of spectral degradation (as implemented by noise-excited vocoding), significant differences in both pupil size and reaction times are still being detected as spectral degradation worsened, suggesting that there is an ongoing increase in the costs of effortful listening (Pals et al., 2013; Winn et al., 2015).

There is, thus, the unsettling possibility that, in the case of the CI user performing exceptionally well within a speech intelligibility assessment, this achievement could be due to unacceptably high levels of LE. Yet another disquieting notion is that there is evidence to suggest that hearing impairment (or, indeed, auditory impoverishment in the case of the CI user) needs only be minimal before LE levels are augmented, and adversely so (McFadden & Pittman, 2008; Rabbitt, 1990; McCoy et al., 2005).

Thus, it appears that the CI user is vulnerable to a vicious cycle of cognitive depletion which then brings with it the potential for severe and pervasive allostatic injury, increasing the risk of disease, exhaustion and even physical and mental breakdown.

1.2. Understanding listening effort

Indeed, using the existing literature already cited, it is possible to extract a series of key factors that may contribute to both the genesis and impact of LE. Currently, these factors exist in the literature in isolation from one another. However, it is theoretically possible that these individual elements can actually be connected together into a more cohesive model of cause and effect. Therefore, a theoretical model of these interconnections is proposed (Figure 1.3). Within this model, every single element represents concepts already discussed in this literature review. The red arrows are intended to indicate where the interaction may be bidirectional. Whenever an arrow is dashed (regardless of colour), it represents occasions where the given factor has a potentially deleterious effect. The black continuous arrows indicate the relationship between the CI user and the FUEL model. The blue continuous arrows highlight where psychosocial consequences may be implicated. The brown continuous arrows identify the factors that may contribute to the CI user's neurophysiology and auditory processing.

It is important to emphasise that the proposed interconnections (in Figure 1.3) are theoretical, with no explicit substantiation available in the current literature related specifically to CI users. However, if these interconnections do indeed exist, they provide a potential pathway demonstrating how LE becomes a health risk for the CI user (especially if LE becomes chronic).

In light of this threat to both physical and mental health, the existing (and widely reported) complaint of LE within the deaf population needs to be taken seriously and be actively addressed within clinical practice and rehabilitation. This, therefore, necessitates accurate measurement of LE (by the audiologist) to assess its true impact.

1.2. Understanding listening effort

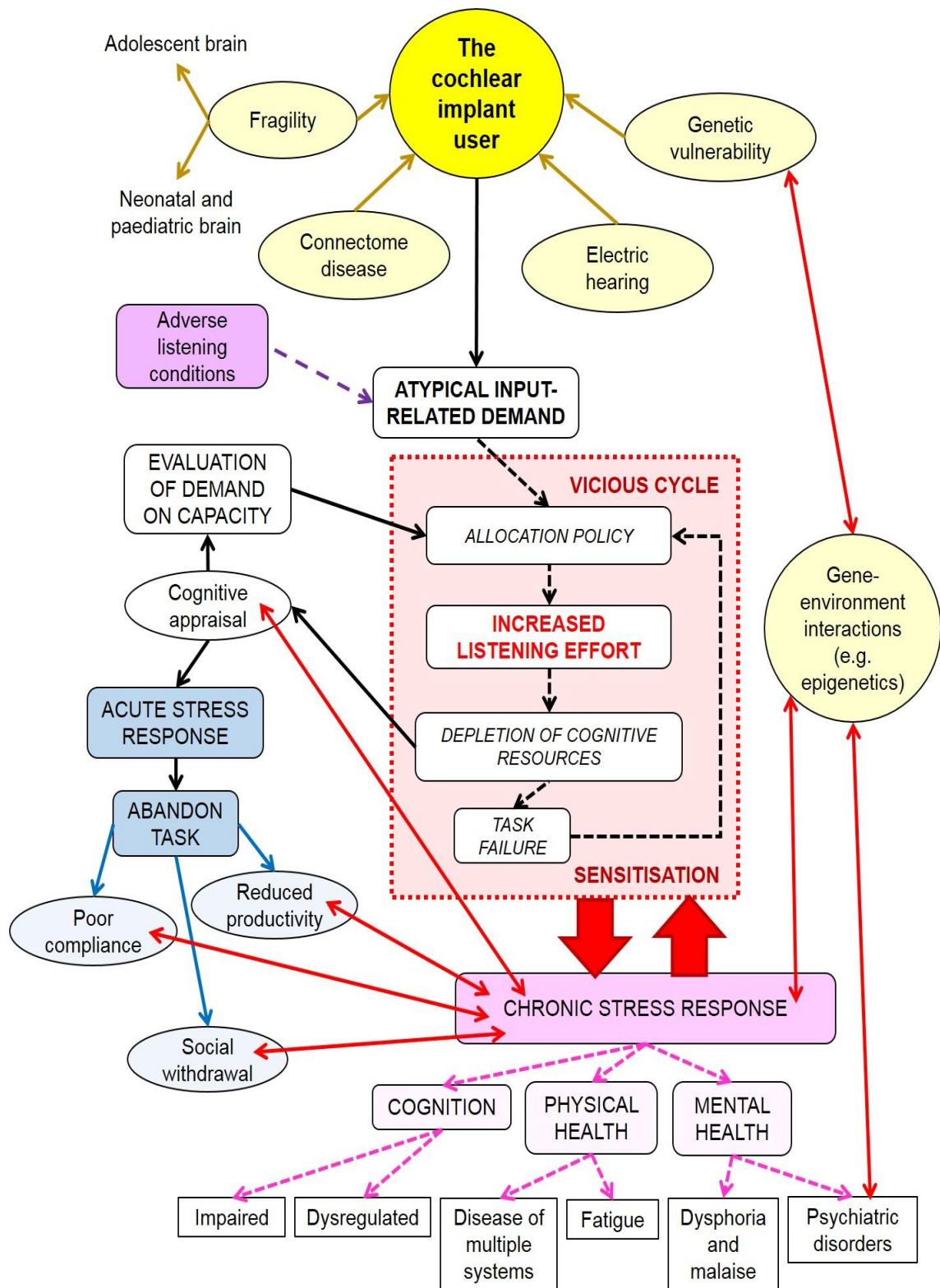


Figure 1.3: Schematic providing a summary of the vicious cycle of cognitive depletion that a deaf individual (and the cochlear implant user) is liable to enter, as well as its interactions with the manifold dimensions associated with the stress response (both acute and chronic); the connectome disease; electric hearing; and any inherent vulnerability and fragility.

1.2.3. How can listening effort be measured?

Currently, there is no clinically standardised way of measuring LE for the hearing impaired, let alone the CI users (Hughes et al., 2017; Pichora-Fuller et al., 2016; McGarrigle et al., 2014). This means that it is difficult to answer the question as to what extent CI users actually experience LE (compared to their normal hearing counterparts). The current literature relating LE directly to the CI user population is sparse. There have been attempts to develop different forms of LE measurement, some of which might be appropriate for use within the CI population. However, the undertaking to quantify and interrogate LE is by no means easy (Lemke & Besser, 2016). There has even been disagreement as to the possibility of LE being measurable at all (McGarrigle et al., 2014). In line with the abundance of working definitions for LE, there has also been a proliferation of putative measures suggesting different proxies for LE (McGarrigle et al., 2014; Lemke & Besser, 2016).

Overall, there are three main approaches that have been attempted in LE measurement: subjective; physiological; and behavioural (McGarrigle et al., 2014; Pichora-Fuller et al., 2016). Examples of LE studies specifically targeting the CI user cohort will be reviewed within the context of these categories of LE measurement, together with the more prevalent studies within the normal hearing population (including those using CI simulations).

1.2.3.1. Subjective measures

Self-report is a subjective measure designed to assess the individual's self-perceived effort, by assuming individuals are able to introspect on their cognitive processes (Paas et al., 2003; McGarrigle et al., 2014). These self-report measures tend to be closed-set questionnaires (Humes & Humes, 2004), such as the Speech, Spatial and Qualities scale (SSQ: Gatehouse & Noble, 2004; Noble & Gatehouse, 2006), or visual analogue rating scales (Pichora-Fuller et al., 2016; Rudner et al., 2012; van Esch et al., 2013). Most rating scales range between the report of "no effort" to that of "maximum effort", and are utilised to assess the self-reported momentary allocation of cognitive capacity and/or perception of task difficulty of the input-related demands of listening, either during or after a set of trials (Pichora-Fuller et al., 2016; Lemke & Besser, 2016; Winn et al., 2015).

1.2. Understanding listening effort

Some rating scales attempt to elucidate the motivation dimension, instead of the demand dimension, by addressing the individual's motivation to complete a task and/or the importance of success (Kramer et al., 2016; Pichora-Fuller et al., 2016).

In the case of questionnaires, the assessment of the LE construct tends to be extracted from an existing multidimensional and multi-item inventory, which encompasses more than LE alone by addressing the nature and extent of listening difficulty (Paas et al., 2003; Pichora-Fuller et al., 2016). These questionnaires tend to elicit a retrospective judgement about the LE involved in everyday listening, whereas judgements in a rating scale are relatively more real-time, being performed either during or immediately after a trial (Rennies et al., 2014; McGarrigle et al., 2014). Some self-report measures target the fatigue element that is often associated with LE, rather than explicitly effort. These too tend to be in the form of a questionnaire, or a scale (Kramer et al., 2006; Nachtegaal et al., 2009; Alhanbali et al., 2017).

Such subjective measures have already revealed a trend of hearing impaired listeners providing higher LE ratings than their normal hearing counterparts (van Esch et al., 2013; Brons et al., 2014). The effects of acclimatisation with assistive technology (such as hearing aids) have been also been investigated using LE-associated ratings, revealing trends of perceived effort becoming significantly less in new hearing aid users after months of acclimatisation (Dawes et al., 2014). Some self-report measures have also been applied within the CI user cohort (but these are sparse), with questionnaires and ratings scales being utilised, as well as the analysis of verbatim transcripts (collected from focus groups) using the principles of constructivist grounded theory methodology (Perreau et al., 2017; Hughes & Galvin, 2013; Hughes et al., 2017).

This class of subjective measurements is particularly appealing as LE assessment, because they are intuitive and straightforward to administer, with first-hand information being immediately obtained and requiring no expertise for interpretation (McGarrigle et al., 2014; Mackersie et al., 2014; Lane & Mackersie, 2015). Indeed, new self-report measures are continuously being innovated. For instance, recently there was a new measure that targets the individual's lowest acceptable performance level (Boothroyd & Schauer, 2001). The aim of this particular self-report estimation is to gauge when a listener was likely to give up listening, with the individual being asked to indicate how long they would be able, and willing, to sustain attention in a given hypothetical scenario with various signal-to-noise ratios (SNRs) of background noise. Not only this, ratings regarding the nature of the SNR are also included (i.e. how loud, how annoying and how distracting the given SNR is).

However, it must not be ignored that self-report measures have their limitations, and some of these stem from the very fact that this approach is by its nature subjective (McGarrigle et al., 2014). After all, it may be the case that there are individual differences in LE “threshold” and in the internal reference scale, i.e. what one individual deems to be effortful (and how quickly effort builds up) may not equate with what another person constitutes as effort (McGarrigle et al., 2014). For example, there is research to suggest that older adults may have the tendency to underestimate their self-perceived LE, with less extreme ratings provided by the elderly compared to younger adults (Larsby et al., 2005).

Another related confound is the notion that there may be differences in interpretation of what LE means, with some individuals perhaps using task performance accuracy or task difficulty as their criterion, rather than the mental exertion per se (Feuerstein, 1992).

Therefore, for some individuals, the changes in the mental exertion, or the mental effort implicated in LE, may become undetectable when using exclusively self-reported measures (McGarrigle et al., 2014).

Yet another issue with subjective ratings, in general, is that the participants’ responses within these measures tend to be provided with some degree of hindsight; thus, the individual may be influenced by their current state of mind at the time of reflection, rather than the state of mind induced by the given testing condition (McGarrigle et al., 2014). Indeed, it has been found that a participant in a relatively positive state of mind might underestimate the extent of their fatigue and vice versa (Deluca, 2005). To exacerbate matters further still, some of these self-report measures claiming to target LE may in fact be too generic, instead tapping more into the overall sensation of chronic stress or fatigue, rather than purely LE itself (Hornsby et al., 2016; Nachtegaal et al., 2009; Pichora-Fuller et al., 2016).

The issue of potential difference in thresholds, and/or interpretation of the meaning in personal criteria, applies to the scales addressing the LE-induced fatigue as well (Hornsby, 2013). Indeed, it also needs to be noted that, despite there being an intuitive link between the effort induced during real-time listening and the fatigue that emerges as a consequence of that effort (with fatigue perhaps being related to long-term LE), there is actually not much empirical support available in the literature for this connection (McGarrigle et al., 2014; Lemke & Besser, 2016).

1.2.3.2. Physiological indices

A class of LE measurement that avoids this subjective element entirely, being instead objective, is the physiological approach. Physiological measures refer to the recording of changes in the nervous system activity during task performance (McGarrigle et al., 2014; Pichora-Fuller et al., 2016; Kahneman, 1973). The assumption underlying these physiological indices is that any cognitively and/or perceptually challenging task will reveal systematic physiological changes that occur during task performance (Piquado et al., 2010; Zekveld et al., 2011; Peelle et al., 2010). Thus, should augmentation in a particular physiological process be exhibited in the more challenging listening condition (relative to a less challenging listening condition), then the physiological change can be attributed to increased LE (McGarrigle et al., 2014; Pichora-Fuller et al., 2016).

These physiological measurements fall into two main categories: measures of the central nervous system (typically activity levels of the brain itself); and measures of the autonomic nervous system (McGarrigle et al., 2014; Pichora-Fuller et al., 2016).

Central nervous system measures in LE research have included functional magnetic resonance imaging (fMRI); magnetoencephalography (MEG); electroencephalography/event-related potentials (EEG/ERPs); and more recently, functional near infra-red spectroscopy (fNIRS) (McGarrigle et al., 2014; Pichora-Fuller et al., 2016; Wiggins et al., 2016; Defenderfer et al., 2017). These techniques do vary in spatial and temporal resolution, thus affecting the nature of information obtained with regards to the timing and location of the brain activity apparently associated with LE (Pichora-Fuller et al., 2016).

The technique of fMRI yields the most precise region-specific localisation of brain activity (Pichora-Fuller et al., 2016). With the assumption that blood oxygen level indexes brain activity, increased neuronal activity (during task performance) conveys an increased metabolic demand for oxygenated blood (Pichora-Fuller et al., 2016). Revelations courtesy of fMRI (with blood oxygen level-dependent contrast imaging to visualise the haemodynamic response) within LE research include the observation that frontal brain regions (in particular) tend to demonstrate an elevated haemodynamic response when listening tasks are challenging (Vaden et al., 2013, 2015). These elevations in blood oxygenation levels, particularly within the cingulate cortex, have been interpreted to reflect a decision-making process, or a cost-benefit analysis, regarding the expected value of potential work required (to optimise performance) relative to the potential value realised from the reward upon successful task completion (Richter, 2016; Pichora-Fuller et al., 2016).

1.2. Understanding listening effort

The fMRI technique has been used to assess the effect of CI simulations on speech processing (generated by noise-excited vocoding being applied to the auditory stimuli) on effortful listening (Wild et al., 2012). Increased blood oxygenation levels were yielded in the left inferior frontal gyrus when participants were required to focus on degraded speech with distractors present, compared to clear speech (Wild et al., 2012). These relative changes in frontal brain activity in the left hemisphere were attributed to the compensatory effort required to cope with, and attend to, the impoverished speech signal (Wild et al., 2012; McGarrigle et al., 2014).

As with LE measurement involving the use of self-report measures, physiological data from actual CI users are scarce. This paucity is not helped by the fact that the CI neuroprosthesis either contraindicates traditional neuroimaging techniques (such as fMRI), or renders execution of them difficult (e.g. with the production of electrical artefacts in EEG measurement) (Wiggins et al., 2016). However, fNIRS offers a potential solution, being completely compatible with the CI neuroprosthesis, as well as the additional advantages of being non-invasive, portable, relatively inexpensive, and requiring a low degree of participant tolerance (Wiggins et al., 2016).

fNIRS has already demonstrated potential as a LE measurement recently within the normal hearing population listening to CI simulations. Changes in neural activity (inferred from the haemodynamic response) have been detected in both the temporal lobe and also the inferior frontal gyrus (Defenderfer et al., 2017; Wijayasiri et al., 2017). These haemodynamic responses have been suggested to be potential neural markers of LE (Wijayasiri et al., 2017).

The technique of ERP, unlike fMRI, has poor spatial resolution (Pichora-Fuller et al., 2016; McGarrigle et al., 2014). Its strength instead is that it offers the most precise temporal information (out of all current neuroimaging techniques that are compatible with the CI device), thus enabling the investigation of time-locked neural activity evoked by stimulus presentation and response (Trembley & Backer, 2016). The measurement of this time-locked neural activity is obtained via recordings of the electric potential fluctuations detected through a series of electrodes placed directly onto the scalp (McGarrigle et al., 2014).

The typical oscillatory neuronal activity (i.e. regular waveform fluctuations) elicited during these scalp recordings tends to have specific temporal and morphological characteristics, with components (or event-related potentials) of the oscillation deemed to reflect certain properties, or aspects, of brain function (Obleser & Kotz, 2011; Bernarding et al., 2012).

1.2. Understanding listening effort

Within the context of LE research, the P3a and also the N1 component have been of interest (Combs & Polich, 2006; Bertoli & Bodmer, 2014, 2016; Obleser & Kotz, 2011; Bernarding et al., 2012). The P3a (evoked by auditory stimuli) has positive amplitude which peaks over the frontal or central electrode sites on the scalp (with peak latency within 250 and 280 milliseconds) (Polich, 2003). The P3a is believed to be particularly associated with brain activity related to orienting of attention (and other involuntary shifts), as well as the processing of stimulus novelty (Polich, 2003).

It has been found that, when the listening difficulty increases (because of decreasing SNR), the amplitude of the P3a becomes reduced and its latency delayed (Trembley & Backer, 2016; Combs & Polich, 2006; Bertoli & Bodmer, 2014, 2016). Accordingly, the P3a has been postulated as a potential index of LE. However, this interpretation needs to be taken with caution, as the changes in morphological and temporal characteristics in the P3a may be reflecting some other correlate (or correlates) of SNR. Indeed, the meaning of the P3a could be even more prosaic, by simply being the response to a change in SNR.

In the case of the N1 component (a negative polarity potential typically occurring 100 millisecond post stimulus onset), CI simulations have been tested (again using noise-excited vocoding) and appear to elicit early peaks and increased amplitude (Obleser & Kotz, 2011). This change in temporal and morphological properties is interpreted as increased effort required to decode the degraded message (Obleser & Kotz, 2011). Phase synchronisation is another property investigated in the N1 component, which relates to how well a given recording for a condition is time-locked to the successive trials (Bernarding et al., 2012). It was revealed that, with simultaneous presentation of syllables (with either similar or different-sounding syllables) in a syllable-detection task, the phase synchronisation of the N1 component systematically increased with the similarity in syllables presented (Bernarding et al., 2012). Since the similar syllable condition is deemed to be the difficult condition, this increased synchronisation is interpreted to correspond to increased LE (Bernarding et al., 2012).

In the related technique of EEG (also of high temporal resolution), the focus is on changes in oscillatory power (Pichora-Fuller et al., 2016). Of particular interest are the alpha and theta responses, which are two of the multiple classes of oscillatory activity found in the brain, as defined by the frequency band of the oscillations (Pichora-Fuller et al., 2016). It has been postulated that this change in alpha power has a functional role in gating local neuronal circuits, using inhibition to control level of activity (Weisz et al., 2011; Sauseng & Klimesch, 2008).

Within the context of LE research, alpha power (with oscillation rates lying between 8 and 13 Hz) tends to be interpreted to reflect change in demand related to the storage and inhibition of information (Pichora-Fuller et al., 2016). In particular, enhanced alpha oscillations are documented to be the neural substrate of cognitive effort (Weisz et al., 2011). It has been shown that alpha power increases when listening in the presence of acoustic degradation, potentially indicating modulation of LE in response to input impoverishment (Obleser et al., 2012, McGarrigle et al., 2014).

The other category of physiological measurements, i.e. of the autonomic nervous system, include the assessment of change in skin conductance, change in pupil size, cardiac responses, and hormonal responses (Kramer et al., 2016).

All these measures ultimately are the output of the summed activity of the parasympathetic and sympathetic nervous systems responding to the level of arousal, which in turn can be influenced by changes in task demand, such as those experienced in everyday listening (Kramer et al., 2016; Pichora-Fuller et al., 2016).

In the case of skin conductance, this response relates to the amount of moisture present on the surface of the skin, as quantified by measuring the skin's capacity to conduct an electrical current (Boucsein, 2012). This moisture is related to the activity of the eccrine sweat glands (principally found on the palm of the hand, or the soles of feet) which are in turn governed by parasympathetic and sympathetic activity reflecting changes in arousal (McGarrigle et al., 2014). Skin conductance became a candidate for LE assessment when it was observed that an increased skin conductance response is elicited when there are increases in listening demand during execution of tasks such as speech repetition or dichotic digit detection (Mackersie & Calderon-Moultrie, 2016; Mackersie & Cones, 2011). This change in skin conductance appears to be reasonably consistent at the individual level too, as well as a group effect (Mackersie & Cones, 2011).

Pupil size in the eye (as assessed by pupillometry) is controlled by muscle activity in the iris, and these muscles too are mediated by parasympathetic and sympathetic activity (Beatty & Lucero-Wagner, 2000; Loewenfeld, 1993). Pupil size has been used to investigate cognitive processing for many years (Janisse, 1977; Kramer et al., 2016; Sirios & Brisson, 2014). As a result, there is ample evidence demonstrating that pupil size can systematically change related to task-evoked load and mental effort during a variety of mental tasks in both normal and clinical populations (Zekveld et al., 2010, 2011; Koelewijn et al., 2012; Kramer et al., 2012).

For instance, peak pupil dilation is believed to index momentary load, whereas the resting pupil diameter before and after stimulus presentation reflects the individual's state of engagement (Pichora-Fuller et al., 2016).

With regards to LE itself, consistent and directed changes in pupil size have been exhibited, revealing pupillometry to be sensitive and highly granular in measurement of the LE response to speech intelligibility, type of background noise, syntactic complexity, lexical manipulation, as well as specific auditory characteristics (Winn et al., 2015; Zekveld et al., 2010; Koelewijn et al., 2012; Kramer et al., 2012; Kuchinsky et al., 2013; Liao et al., 2015). The imposition of hearing impairment too has been found to elicit pupil dilation (Kramer et al., 1997; Zekveld et al., 2011).

Pupillometry has been conducted with both CI simulation and actual CI users themselves (Winn et al., 2015; Steel et al., 2015). In one example of a CI simulation study (Winn et al., 2015), two types of vocoding were tested: a noise-excited vocoder where the number of spectral channels used were varied; and a vocoder involving peak-picking channel selection and variable synthesis slopes (designed to mimic the front-end processing and spread of neural excitation typical of the CI). Both revealed pupil dilation growing with each successive degradation of spectral resolution (Winn et al., 2015).

In the case of the CI population, within a binaural fusion study (comparing bilaterally implanted children to normal hearing children), pupil diameter is concomitant with the child's response according to binaural fusion in both cohorts, with pupil size increasing with worsening binaural fusion (Steel et al., 2015). Furthermore, it was found that children have difficulty fusing input from their bilateral implants to perceive one sound, which is then believed to cost them increased LE as corroborated by the increased pupil size (Steel et al., 2015).

In the case of the cardiac response, two aspects have been of particular interest. The first is the heart-rate variability (HRV) (Pichora-Fuller et al., 2016; Mackersie & Calderon-Moultrie, 2016). Measures of HRV quantify the amount of heart variation in both time and frequency domains. It has been revealed that there is an overall reduction in HRV with increased listening demand, with different types of HRV measurements consistently showing this trend (such as standard deviations of interbeat intervals, spectral analysis of variation in interbeat intervals, and square root of the mean squared differences between normal beats) (Pichora-Fuller et al., 2016; Mackersie & Calderon-Moultrie, 2016).

1.2. Understanding listening effort

The second aspect of cardiac response that has also been of research interest, especially for those studying motivational intensity theory, is the pre-ejection period (PEP), the time interval between the beginning excitation of the left heart ventricle and the opening of the aortic valve (Brehm & Self, 1989; Wright et al., 1996; Richter, 2016; Chatelain & Gendolla, 2015; Richter et al., 2016). PEP directly indicates the myocardial contraction force (whereby the stronger the contraction, the shorter the PEP) and is determined by sympathetic activity (Pichora-Fuller et al., 2016; Richter, 2016).

Finally, measures of the hormonal response of the autonomic nervous system involve the use of endocrine biomarkers (such as cortisol and chromogranin A), which can be obtained via salivary samples (Pichora-Fuller et al., 2016; Kramer et al., 2016).

There are relatively few studies that have measured these biomarkers in the context of LE and these have revealed that cortisol levels tended to be relatively higher in the hearing impaired, when compared to normal hearing controls (Hicks & Tharpe, 2002; Jahncke & Halin, 2012; Kramer et al., 2016). Reduced secretion of these hormones have also been postulated to index fatigue (McGarrigle et al., 2014, Hicks & Tharpe, 2002).

Overall, physiological approaches are a particularly convenient method of conducting LE measurement because all available measures are non-invasive and can be executed without disrupting the listening task, meaning that there is little interference of the attempted LE measurement on the listening task (McGarrigle et al., 2014).

However, it is important to recognise that all physiological indicators are essentially indirect in their capture of LE, meaning there is a need for well-controlled and well-designed experiments before any inference can be drawn (Winn et al., 2015). Even then, all these indices are still only correlates of LE, and correlation does not automatically assume causality (Pichora-Fuller et al., 2016). Indeed, their precise relationship with LE is still yet to be clarified (Kramer et al., 2016; Pichora-Fuller et al., 2016).

Researchers have begun to try to render the relationship of these indices with LE more concrete by exploring the associations present between the physiological measures themselves, yielding both reassuring and concerning conclusions (Kramer et al., 2016; Pichora-Fuller et al., 2016). The first concern is that, within studies that have concurrently obtained multiple physiological indices at the same time, it appears that the different indicators are not equally affected by the listening demand, irrespective of whether this demand has been imposed by either the experimental testing conditions, or by a hearing impairment (Kramer et al., 2016).

Secondly, the exclusivity of the autonomic activity in reflecting LE has been under question too. Indeed, in the case of the pupil response, it has been suggested that the increase in pupil size reflects overall cognitive duress, rather than the specific LE of the listening task at hand (Zekveld et al., 2014a; Koelewijn et al., 2012). Another suggestion is that pupil size is more an indicator of brain activity level (i.e. whether a more extensive brain network is being utilised, or whether brain activity has intensified). However, these alterations in brain activity cannot be automatically assumed to be specifically related to LE (Koelewijn et al., 2012; Grady, 2012).

An alternative account proposed is that different aspects of the pupil response are associated with different aspects of cognitive processing (Pichora-Fuller et al., 2016). After all, pupil size has long been believed to be modulated by attention, memory, and even cognitive ability (Sirios & Brisson, 2014; Zekveld et al., 2010; Koelewijn et al., 2012; Karatekin, 2004).

Motivation has also been implicated in the pupil size (Kahneman et al., 1968; Varazzani et al., 2015; Hopstaken et al., 2015). Indeed, it was found that, when sufficient rewards were presented to a fatigued individual, the pupil dilation response could be restored (having previously become constricted), suggesting increased motivational harmony for task performance in light of reward attainment and/or success importance (Hopstaken et al., 2015; Richter, 2016; Matthen, 2016). Pupil constriction elicited by changes in environmental light, on the other hand, is believed to reflect pure parasympathetic activity (Kramer et al., 2016).

To further support the importance of motivation, the pupil size has also been demonstrated to be impacted by the activity of the locus coeruleus (Murphy et al., 2011). The locus coeruleus is a subcortical structure which is the sole source of norepinephrine (sometimes called noradrenaline), with dense cortical and subcortical connections throughout the entire brain, meaning that it is implicated in cognition and arousal (both vital for determining motivation), as well as the sleep-wake cycle (Aston-Jones & Cohen, 2005; Sirios & Brisson, 2014). The locus coeruleus has been shown to be itself modulated by processes such as attention, stress and memory (Laeng et al., 2012; Varazzani et al., 2015; Murphy et al., 2011).

The same level of uncertainty applies to the skin conductance response, with this measure also being linked with automatic attention (e.g. orienting), motivation, and also emotional reactivity (Kahneman, 1973; Andreassi, 2007; Bourscein, 2012; Mackersie & Calderon-Moultrie, 2016). Additionally, the findings related to the endocrine biomarkers have too been confounded by other processes potentially having a role, such as that of stress and arousal (Kramer et al., 2016).

Indeed, the association of hormone secretion with LE has been inconsistent and remains controversial; for example, it is not clear whether these measures are simply insensitive, or whether stress has to accumulate sufficiently to be detected in biological samples (McGarrigle et al., 2014; Hicks & Tharpe, 2002; Kramer et al., 2016).

While the model of FUEL (the Framework for Understanding Effortful Listening: Pichora-Fuller et al., 2016) does implicate these extra variables (e.g. motivation and arousal) to be part of the LE construct, complication still remains in data interpretation and it still cannot be guaranteed that LE is actually being measured (however indirectly), because these additional processes have the ability to emerge of their own accord (McGarrigle et al., 2014).

A further confound affecting the reliability and sensitivity of physiological measurement in general is the fact that there is a great deal of background physiological activity that naturally arises from the moment-to-moment operation of the human body and brain (Winn et al., 2015). Thus, the activation of interest (i.e. that specifically evoked by LE) is likely to be hidden or obscured by this background noise (Winn et al., 2015). This renders many of the physiological techniques (such as those involving pupillometry) especially intensive in data processing, because of the need to clean, extract and/or normalise the fluctuations of interest (Piquado et al., 2010). Even then, it cannot be certain that the complex mathematical algorithms have actually succeeded in this extraction (Winn et al., 2015).

To obscure interpretation of physiological data even more, there are often inconsistencies in yielded data that cannot always be accounted for by differences in experimental design (Zekveld et al., 2014a, b). For instance, in apparent response to especially challenging listening conditions, such as that of very low speech intelligibility, the pupil size has been reported to increase (Cabestrero, 2009), whilst other studies report the pupil response to decrease (Zekveld & Kramer, 2014; Granholm et al., 1996). Similar inconsistencies in pupil response have been found in the hearing impaired population too (Kramer et al., 2016).

In the instances of the pupil size constricting (instead of dilating) in the face of acoustic challenge, a suggested explanation has been that the brain has reached cognitive overload (Zekveld & Kramer, 2014). Cognitive overload within this context is argued to be the complete overwhelming of the finite capacity of cognitive resources, meaning that the brain disengages from the task at hand, thus engendering no further demand on the brain (and, therefore, body), hence the diminishing of pupil size (Zekveld & Kramer, 2014). This is supported by evidence that, when pupillometry is combined with EEG, pupil size reduces with task disengagement (Hopstaken et al., 2015).

1.2. Understanding listening effort

Another potential account for the inconsistencies found in this physiological data is that there could be individual differences confounding measurement (McGarrigle et al., 2014). Indeed, there is a wealth of literature demonstrating individual differences in cognitive ability, and it has been postulated that perhaps the pupil size is actually measuring, or at least is confounded by, working memory capacity (Rönnberg, 2014; Heitz et al., 2008; Rudner et al., 2012).

There have been studies (utilising the Reading Span task) observing that smaller pupil sizes are found in individuals with high working memory capacity (Heitz et al., 2008). Within this context, it has been postulated that a high working memory capacity could help to compensate for any mismatch of the acoustic signal to the stored internal representations (that arise due to signal degradation imposed by deafness), which would then increase the ease of understanding and thus reduce the required LE (Rönnberg, 2014; Foo et al., 2007; Lunner, 2003; Rudner et al., 2009).

Indeed, there is data from subjective ratings consistent with this notion, with a demonstrated association of high working memory capacity with lower perceived effort (Rudner et al., 2012).

1.2.3.3. Behavioural assessments

Behavioural assessment is another objective technique that has been used in LE measurement (McGarrigle et al., 2014). These measures can be divided into two types: single-task and multi-tasking paradigms (McGarrigle et al., 2014). A corollary technique, which is not as frequently used, is the behavioural measurement of mental fatigue following exerted LE (McGarrigle et al., 2014; Deluca, 2005).

At the behavioural level, mental fatigue is characterised as the slowing, or decline, of cognitive functions following concerted or prolonged mental task performance (Deluca 2005). This is typically detected using a vigilance task, in which the participant sustains their attention for a prolonged duration (McGarrigle et al., 2014). Within the context of LE, it has been demonstrated that there is a behavioural slowing in visual response time over the duration of a test block of a word recognition task when executed by hearing impaired individuals (Hornsby, 2013). In addition, not only was there this apparent LE-induced cognitive fatigue, but also this fatigue seemed less marked in aided versus unaided listening conditions as revealed by within-participant comparisons (Hornsby, 2013).

1.2. Understanding listening effort

In the case of the single-task paradigm as applied in LE research, it is a straightforward assessment to implement, in which the participants typically are required to respond to speech stimuli in a variety of listening conditions, either by verbally identifying the stimulus, or by pressing a response button (Gatehouse & Gordon, 1990; Houben et al., 2013).

The time taken for the response to be produced (be it verbal, the pressing of a button, or variations thereupon) is the variable of interest, with it being stipulated that the speed of response may provide additional information regarding the LE related to speech perception (Gatehouse & Gordon, 1990; Houben et al., 2013).

Indeed, in a study with normal hearing participants performing digit triplet identification in varying levels of stationary noise, reaction times were found to be significantly slower in lower signal-to-noise ratios (SNRs) in spite of optimal intelligibility (Houben et al., 2013).

This increased latency in reaction was taken to reflect the increased LE necessitated to understand the digit triplet at the lower SNR (Houben et al., 2013; McGarrigle et al., 2014).

There is another more novel format that the single-task paradigm can take with regards to LE measurement, which is proposed to be sensitive to the fluctuations of LE over a period of more sustained time: the monitoring of uninterrupted speech (e.g. the Glasgow Monitoring of Uninterrupted Speech Task) (MacPherson & Akeroyd, 2013). Instead of presenting short and interrupted sentences, the participant is presented with continuous speech that they are required to monitor over several minutes, whilst simultaneously identifying any word substitutions in a written transcript. Response times in this scenario are believed to correspond to speech processing rate and also LE (MacPherson & Akeroyd, 2013).

With regards to the multi-tasking category of behavioural assessment, the dual-task paradigm is the most common and frequently applied behavioural technique in LE research (Howard et al., 2010; Gosselin & Gagné, 2011a, b; Desjardins & Doherty, 2013). The dual-task paradigm requires the participant to perform a primary and secondary task simultaneously, with the primary task typically involving sentence recognition (McGarrigle et al., 2014; Pichora-Fuller et al., 2016). Face validity for this measure is generally considered to be sound; after all, the execution of multiple simultaneous mental operations pervades everyday life, especially when it comes to speech processing (Howard et al., 2010; Ophir et al., 2009; Wallis, 2006; Watson & Strayer, 2010).

1.2. Understanding listening effort

These dual-task paradigms are based on the same assumption that FUEL is founded upon: i.e. the brain is of finite cognitive capacity (Kahneman, 1973; Broadbent, 1971). As already mentioned, cognitive load can arise from, or be induced by, an assortment of input-related demands, such as linguistic complexity (Lewis et al., 2006) or adverse listening circumstances (Mattys et al., 2012).

In the context of the dual-task paradigm, cognitive load is deliberately manipulated and/or induced by the use of task competition for cognitive resource (Mattys et al., 2012). Should there be an increase in cognitive load for the primary task, there will be a corresponding deterioration in secondary task performance as more cognitive resources are diverted to support execution of the primary task (i.e. at the cost of the secondary task) (Edwards, 2016; Paas et al. 2003). This deterioration in secondary task performance is then inferred to reflect an increase in LE (Downs, 1982; Gosselin & Gagné, 2011a, b; Paas et al., 2003).

The nature of the mental operations, or cognitive domains, utilised within these dual-task paradigms can be argued to essentially fall into three categories: attention; working memory; and processing speed (Pichora-Fuller et al., 2016; McGarrigle et al., 2014). These three cognitive domains are assumed to be interrelated (Pichora-Fuller et al., 2016; Wingfield, 2016; McGarrigle et al., 2014). Indeed, across these cases, the terms of “cognitive resources”; “processing resources”; “attentional resources”; and “resources” have been used on an interchangeable basis, because all are assumed to essentially represent the same construct, i.e. LE (McGarrigle et al., 2014).

Attention is thought to broadly be involved in the allocation policy of the limited cognitive resources, including the selection and maintenance of information during the performance of a single activity (i.e. selective attention), and also during multiple activities (i.e. divided attention) (Pichora-Fuller et al., 2016). Thus, in the case of the versions of the dual-task paradigms utilising the attention component, it is believed that it is divided attention being manipulated for examination of LE (Phillips, 2016; Best et al., 2010). Different secondary tasks have been used to accompany the speech-related primary task (Gosselin & Gagné, 2011a, b). For instance, in addition to the popular visual secondary task, tactile pattern recognition tasks and even a driving vehicle simulator have been utilised (Gosselin & Gagné, 2011a, b; Hicks & Tharpe, 2002; Fraser et al., 2010; Wu et al., 2016).

Working memory capacity, on the other hand, is believed to be the use of limited cognitive capacity in retaining information in conscious awareness when the information is no longer present in the environment (Wingfield, 2016). This then enables the information to be processed, manipulated and used to guide behaviour.

1.2. Understanding listening effort

This includes complex activities such as language comprehension, or indeed listening whilst multi-tasking (as occurs in the working memory manifestation of the dual-task paradigm: Pichora-Fuller et al., 2016, Postle, 2006; Wingfield, 2016).

There is a wide variety of secondary tasks utilised within the working memory dual-task paradigms (Feuerstein, 1992; Hornsby, 2013; Hornsby et al., 2014). Recall is a popular candidate, whereby the ability of the participant to remember and correctly retrieve the target encoded item is tested (Lunner et al., 2016). An example of a concurrent memory task is listening to various talkers, which has been shown to increase LE (when compared to a single talker), even if these various talkers were presented sequentially (Nusbaum & Morin, 1992).

Processing speed is argued to be symptomatic of (or at least related to) residual cognitive capacity, whereby the speed of processing slows as the amount of capacity consumed by task performance increases; if available capacity is exceeded by task demand, processing speed is believed to be sacrificed to minimise, or remove, erroneous responding (Lemke & Besser, 2016; Pichora-Fuller et al., 2016). Dual-task paradigms using probe reaction time are intended to test processing speed and/or attention, thus forming the final assembly of dual-task paradigms (Downs, 1982; Desjardins & Doherty, 2013).

Examples of this index within LE research include measuring reaction time in the performance of a non-auditory task (such as simple versus choice reaction time to visual stimuli) and an auditory task, to examine domain-general variance associated with the overall processing speed, as well as the domain-specific variance associated with auditory processing speed (Deary et al., 1989; Deary, 1994; Philips, 2016). Even when accuracy of the auditory task is at ceiling, this index of processing speed has been informative in revealing the effects of acoustic distortions and semantic context on listening (Goy et al., 2013).

With regards to the use of the dual-task paradigm with the hearing impaired and CI user populations, hearing loss too has been shown to interfere with memory within a recall dual-task (e.g. interference with visual memory: Rakerd et al., 1996). The recall dual-task paradigms have also been used in the evaluation of the effects of hearing loss and hearing aid use on LE levels (Lunner et al., 2016).

CI simulations have also been pursued in order to identify how the impoverishment of the speech signal (that is typically produced by the CI neuroprosthesis) may lead to differential resource allocation according to how effortful listening has become (Pals et al., 2013).

For instance, one study incorporated a reaction time dual-task and found that different levels of spectral degradation (as implemented by differing numbers of spectral channels in noise-excited vocoding) impacted on reaction time (Pals et al., 2013). A systematic decline in reaction time (i.e. improvement in performance) was observed as the number of channels were increased from 2 to 8 channels (in 2-channel steps), suggesting that signals with better resolution required less effort to be understood (Pals et al., 2013). Visual reaction time was also found to be reduced in a CI simulation study comparing comprehension of a native language versus a second language (Ganesh et al., 2011).

As for actual CI users, available literature applying the dual-task paradigm with this cohort is as rare as for the physiological or subjective methods of LE measurement. One example entailed the use of a visual Stroop reaction time task (within a dual-task paradigm) in testing of normal hearing controls and unilateral, bilateral, and short-electrode (with bilateral residual hearing) CI users (Perreau et al., 2017).

It was revealed that there were significant differences in reaction times only between the normal hearing controls and the overall CI user groups across the different signal-to-noise ratios tested (with normal hearing controls eliciting relatively quicker reaction times), but no differences were detected between the different types of CI users themselves (Perreau et al., 2017). Another study compared bilateral CI users and unilateral CI users to normal hearing controls by using reaction time of a visual shape-matching task within the dual-task, with bilateral CI users producing equivalent reaction times to normal hearing controls (Hughes & Galvin, 2013).

Before concluding on the viability of these behavioural measurements for the objective of quantifying LE, the theoretical assumptions underlying these different types of assessments do need to be considered and any alternative hypotheses identified.

In the case of the single-task paradigm, the underlying assumption (i.e. that increased LE would lead to increased latency in response) is plausible due to the notion of challenged cognitive capacity slowing processing (McGarrigle et al., 2014). However, it is not the only possible hypothesis; indeed, it is also plausible that increased LE would actually lead to decreased response time due to more focused attention (McGarrigle et al., 2014). This relationship between the required LE to understand the stimulus and the timing of the response to the stimulus is unclear, thus forming a fundamental limitation for this category of behavioural measurement.

Then there are the cardinal principles of the dual-task paradigm to contemplate. Firstly, for the dual-task paradigm to succeed in measuring fluctuation in LE, it mandates that the individual's entire resource capacity is fully utilised for both tasks, whereby any residual capacity (after resource allocation is executed for the primary task) is designated wholly to the secondary task (Paas et al., 2003). However, it is plausible that there may still be residual capacity not used by either the primary or secondary task (McGarrigle et al., 2014). Indeed, there is currently no independent way of measuring the actual quantity of cognitive resources assigned to each of the two tasks in the dual-task, let alone whether there is any residual capacity in the first place (McGarrigle et al., 2014; Styles, 2006).

It may even be the case that the performance of the secondary task might actually interfere with the processing and/or execution of the primary task (even though it is at lower priority), especially if the primary task is complex (Paas et al., 2003). The level of interference might even reach the stage where the individual decides to focus exclusively on the secondary task (McFadden & Pittman, 2008). Pitching difficulty level thus requires care, with the need for a dynamic range in detectable outcomes to enable sensitivity (Paas et al., 2003).

There is also the supposition that all resource allocation is under conscious control (McGarrigle et al., 2014). Unfortunately, however, it is more than credible that an individual may decide to allocate most of their resources to a relatively more novel task (which is not necessarily the primary task); not only this, the decision could even be unconscious (McGarrigle et al., 2014). For instance, it has been found that children tend to demonstrate a task bias whereby they unconsciously prioritise the more novel task, even if it is against instruction (Choi et al., 2008).

In addition, there is also the potential for contamination caused by individual differences (Watson & Strayer, 2010). Indeed, the phenomenon of "supertaskers" (individuals who possess outstanding multi-tasking ability) has been testified in the literature and their existence cannot be attributed to be a statistical fluke (Watson & Strayer, 2010).

Within a driving simulation experiment that entailed the dual-task paradigm of performing an auditory operation span task at the same time as driving within a simulated environment, it was revealed that (within a cohort of 200 participants) there were participants whose multi-tasking was decidedly superior (Watson & Strayer, 2010). Monte Carlo simulations were also run within this experiment, and they appeared to confirm that the incidence of these supertaskers was above chance level.

1.2. Understanding listening effort

Indeed, this is more evidence in support of the existence of individual differences within cognition, with attention and cognitive control being specified to be of particular relevance to these supertaskers (Watson & Strayer, 2010).

1.2.3.4. Feasibility of LE measurement in clinical environments

Despite there being a healthy bounty of investigative tools for LE, each have their own strength and weaknesses, as well as concomitant complications related to their theoretical underpinnings (McGarrigle et al., 2014; Pichora-Fuller et al., 2016). Furthermore, many of these LE measures (whilst acceptable research tools) are not yet ready for widespread application within the clinical audiological setting in the near future (McGarrigle et al., 2014).

Perhaps, ultimately, the choice of testing method (to pursue for further development as a clinical test) should be directed by the constraints of a routine audiological appointment (i.e. typically attended by a CI user, in order to fine-tune speech perception thresholds and overall sensitivity of the CI device).

This focus on audiological appointments is particularly important because the CI user will be never be discharged from audiological care (becoming a permanent part of the patient load), and will be required to attend appointments for the entire duration of CI device ownership. This is due to the fact that ongoing monitoring and care is needed to ensure that all electrodes are still functioning as the internal device ages. These appointments also enable necessary upgrades as technological advances are made in not only the external device design, but also the processing strategies (used within the frequency-electrode map) and noise-reduction algorithms.

A key constraint to consider (when choosing a LE test) is the limited time available for each audiological appointment. The audiologist already has to execute critical testing regarding speech perception performance (in both quiet and noise) and also check user comfort (in different levels of noise). This means that any additional testing (i.e. required for LE measurement) must be accommodated within the same time span, with no sacrifice of the existing mandatory testing (for detail regarding protocol guidelines for CI audiological appointments, see Appendix 1.1).

Other pragmatic issues involved in introducing a new test into the clinical environment are that the additional equipment and technology requirements are inexpensive and require minimal training for both implementation and data interpretation.

In addition, the test needs to be straightforward for the patient to execute, in order to avoid any confusion or even distress. It is also essential that the test is sufficiently sensitive and reliable to be appropriate for individual testing on a repeated basis. If all of these requirements are satisfied, it may become feasible to include LE measurement as part of standardised audiological assessment for CI users.

Bearing these requisites in mind, especially regarding time constraints, this immediately suggests that a “two-for-one” approach might be a solution, i.e. where two types of test results are yielded by a single procedure. Accordingly, the dual-task paradigm becomes a relevant option, especially if the primary task becomes the speech intelligibility test (e.g. the standardised sentence lists presented at different SNRs). The secondary task (which is conducted simultaneously to the primary task) could then produce a score to glean the index of LE. Another benefit of the dual-task paradigm is that it provides an objective measure of LE.

An advantage of this LE testing for the clinician is that it provides another metric in assessing the current efficacy of the CI device, regarding both its accuracy and sensitivity (in terms of speech intelligibility), and also particularly the cognitive load (incurred by the required auditory processing). This is especially important because, in order to gain maximum benefit from the CI technology, the CI user not only needs to hear well (i.e. achieving optimum speech perception), they also need to hear healthily (i.e. avoiding any allostatic injury triggered by excessive LE).

Therefore, in conclusion, it appears that the dual-task paradigm would be the most tenable methodology to pursue in the quest for a feasible clinical test of LE.

1.3. Statement of aim (of thesis)

In light of the potential perniciousness of LE for the CI population, and in the absence of any clinical assessment of this construct, the aim of this thesis is to evaluate the feasibility of using the dual-task paradigm (i.e. a behavioural approach) as a framework for developing a potential test of LE in CI users.

1.4. Outline for remainder of thesis

This thesis chronicles nine experiments (Figure 1.4): a pilot study, four small-scale studies (with a maximum sample size of 10 participants) and four larger-scale studies (each involving 25-30 participants).

Chapter 2 (in Part 2) archives the pilot study, which entails the first attempt of developing a behavioural dual-task paradigm with both CI users and normal hearing participants (with the normal hearing controls being tested in both normal listening conditions and CI simulations). The pilot study also attempts the use of across-methods triangulation involving pupillometry (a physiological measure) and subjective ratings. This dual-task paradigm utilises auditory recall and visual recall as the primary and secondary tasks respectively.

Chapters 3 and 4 (in Part 3) contain all the experiments involving normal hearing participants in normal listening conditions (i.e. Experiments 1, 2, 3 and 4). Experiments 1 and 2 (in Chapter 3) involve the systematic refinement of both the primary and secondary tasks within the dual-task paradigm, incorporating a visual digit stream (that is presented rapidly) as well as more complex auditory sentence stimuli.

Experiment 3 (in Chapter 3) then tests this refined digit stream dual-task in a large cohort of normal hearing participants in an attempt to establish viability for this LE measurement. Experiment 4 (in Chapter 4) investigates how introducing adaptive tracking into the refined dual-task paradigm affects its sensitivity for LE measurement.

Chapters 5 and 6 (in Part 4) document the testing of the digit stream dual-task with two different participant cohorts: normal hearing participants listening to CI simulations; and actual CI users. Chapter 5 explains the modification of the digit stream dual-task to enable the integration of CI simulations (i.e. Experiment 5 and 6). Chapter 5 also includes the testing of the dual-task within the CI simulations context (Experiment 7).

Chapter 6 reports the final experiment of this thesis (i.e. Experiment 8), which involves 25 CI users.

This thesis concludes with Chapter 7 (in Part 5), which summarises the overall findings of all nine experiments and discusses the implications and potential direction for future research in this field.

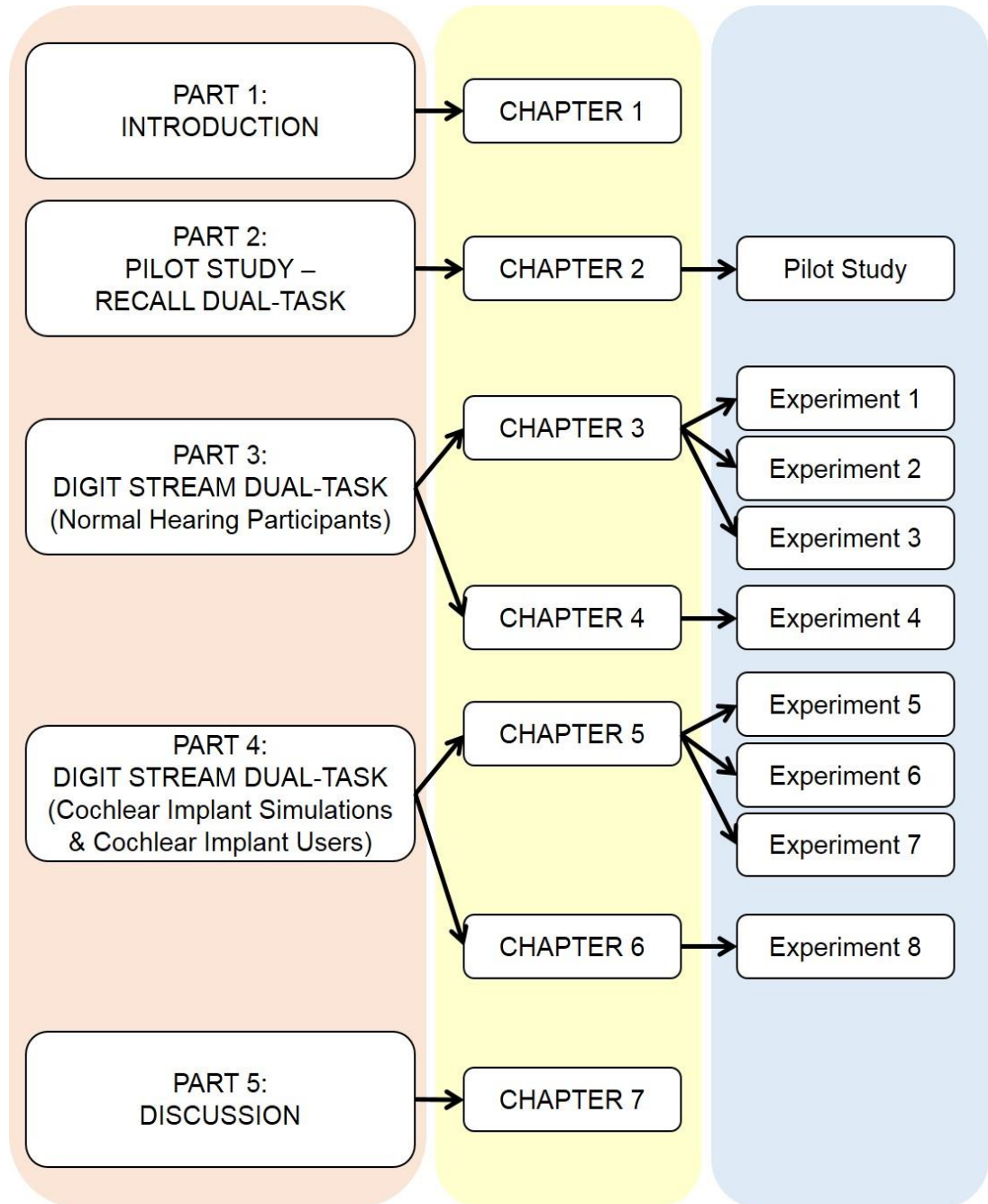


Figure 1.4: Schematic providing a summary of the structure and content of this thesis.

PART TWO:

PILOT STUDY –

RECALL DUAL-TASK

CHAPTER TWO:

Pilot Study testing a Recall Dual-Task

2.1. Introduction

As concluded in the literature review in Chapter 1, the behavioural approach to LE measurement appears to show the most promise in developing a flexible clinical test of LE for CI users. In particular, the dual-task paradigm offers an appealing framework, because of the possibility for simultaneous testing of speech intelligibility alongside LE measurement.

In this chapter, a Pilot Study will be undertaken, whereby a novel version of a recall dual-task paradigm will be trialled with a small group of CI users and normal hearing controls (in both normal listening conditions and CI simulations). An initial small-scale pilot is necessary because there is no precedent for this particular recall dual-task paradigm. Thus, there is a need to collect baseline data before a larger-scale study is pursued.

In order to glean as much information about the potential viability of this recall dual-task, across-methods triangulation will also be attempted. This will be obtained via the concurrent collection of two other indices associated with LE: pupil size (of the eye) and subjective ratings (i.e. the participant's self-report of perceived effort).

Furthermore, to clarify data interpretation as far as possible, the neurocognitive profile of the participants will be assayed, with the objective of achieving reasonable similarity between the two participant groups (i.e. CI users and normal hearing controls). Thus, there will be testing of intelligence, cognitive ability, executive function and also speech-in-noise performance.

2.2. Fundamental principles of the dual-task paradigm

The premise of the dual-task paradigm is that secondary task performance (within the dual-tasking) becomes an index of mental effort (Phillips, 2016; Pichora-Fuller et al., 2016). This is because of two cardinal tenets of cognition:

- Cognitive resources are required for task performance;
- There is a finite capacity to the amount of cognitive resources available.

2.2. Fundamental principles of the dual-task

In line with these tenets, it is posited that, when there is the simultaneous performance of two tasks, the primary task will use up a certain percentage of the available cognitive resources (presumed to be the majority), leaving the remainder to the secondary task, which is lower in priority (Kahneman 1973; Phillips, 2016; Styles, 2006).

Thus, if secondary task performance achieved in the dual-task condition (i.e. when both primary and secondary tasks are executed) is compared to what is achieved in the single-task condition (i.e. when only the secondary task is executed), any deterioration in secondary task performance when dual-tasking indicates how much capacity is being diverted from the secondary task to enable primary task performance (Edwards, 2016; Pichora-Fuller et al., 2016).

It is also stipulated that, should the primary task become more challenging, more cognitive resources will be re-allocated to maintain and/or improve primary task performance in the face of this challenge, to the cost of the secondary task (Phillips, 2016; Pichora-Fuller et al., 2016). Therefore, if secondary task performance achieved when dual-tasking in the presence of this challenge is compared to the secondary task performance when dual-tasking without this challenge, any deterioration in secondary task performance indexes the additional mental effort exerted to address the challenge (Phillips, 2016; Picou et al., 2013; Gosselin & Gagné, 2010).

Furthermore, in order to tap specifically into LE (i.e. a type of mental effort), it is required that the primary task of the dual-task paradigm is a listening task (Pichora-Fuller et al., 2016; Lemke & Besser, 2016). Accordingly, for the purposes of this Pilot Study, the primary task will be a closed-set speech intelligibility task. In addition, the choice of challenge (in order to increase task difficulty for the primary task) will be a masker consisting of speech-shaped noise. This choice of auditory mask is in line with the popularity of using acoustic challenge in order to manipulate task difficulty within LE research (e.g. Howard et al., 2010; Gosselin & Gagné, 2010; Sarampalis et al., 2009).

A particular strength in using noise as a form of acoustic challenge is that it is ecologically valid; indeed, speech is seldom heard in quiet in typical everyday listening conditions. Not only this, speech-shaped noise (whereby its long-term average spectrum is similar to speech) is also a common mask for speech perception research, and is generally acknowledged to be particularly efficacious in providing a baseline in speech intelligibility assessment (Nelson et al., 2003; Qin & Oxenham, 2003; Miller, 1947).

2.2. Fundamental principles of the dual-task

In an attempt to avoid any conflict in auditory processing between the primary and secondary task of the dual-task paradigm, a different sensory modality will deliberately be used for the secondary task in the Pilot Study: i.e. vision, via the use of a visual recall task. This concept of exploiting an alternative sensory modality (to that required by the primary task) within a dual-task paradigm is not unprecedented, with vision being employed with apparent success in prior LE research (e.g. Downs, 1982; Pals et al., 2013).

Thus, within the context of the Pilot Study, a score of LE can be calculated by subtracting visual recall accuracy when dual-tasking in speech-shaped noise from the visual recall accuracy achieved when dual-tasking in quiet (Figure 2.1). A positive score indicates an increase in LE.

This fundamental assumption that changes in secondary task performance reflect changes in LE is only valid as long as the primary task is prioritised over the secondary task at all times (Paas et al., 2003). However, it may be the case that the execution of the secondary task could interfere with the processing and/or performance of the auditory task, especially if the primary task is complex (Paas et al., 2003). Not only this, the level of interference could even be such that secondary task performance completely dominates primary task performance (McFadden & Pittman, 2008). Potential reasons for this secondary task dominance are manifold, such as the secondary task being relatively more novel, meaning it acts as a distractor (Choi et al., 2008; McGarrigle et al., 2014). Accordingly, there needs to be a checking mechanism to verify that appropriate resource allocation has indeed occurred during testing.

One possible method of checking is to ascertain whether primary task performance changes when dual-tasking, compared to when it is performed on its own. It should be the case that, regardless of the testing condition (e.g. the presence, or absence, of acoustic challenge), the addition of a secondary task does not cause deterioration in primary task performance, i.e. primary task performance stays stable. If deterioration in the primary task does occur in the dual-task condition, this may mean that resources are incorrectly being diverted toward the secondary task.

Therefore, whenever a LE calculation is conducted within this thesis, primary task stability will also be computed to assess the likelihood of appropriate resource allocation having occurred (and, therefore, validity of LE calculation). Within the context of the Pilot Study, primary task stability is derived by subtracting the auditory accuracy achieved when performing the primary task on its own (in noise) from the auditory accuracy achieved when dual-tasking (in noise). A score of zero (or above) indicates primary task stability.

2.2. Fundamental principles of the dual-task

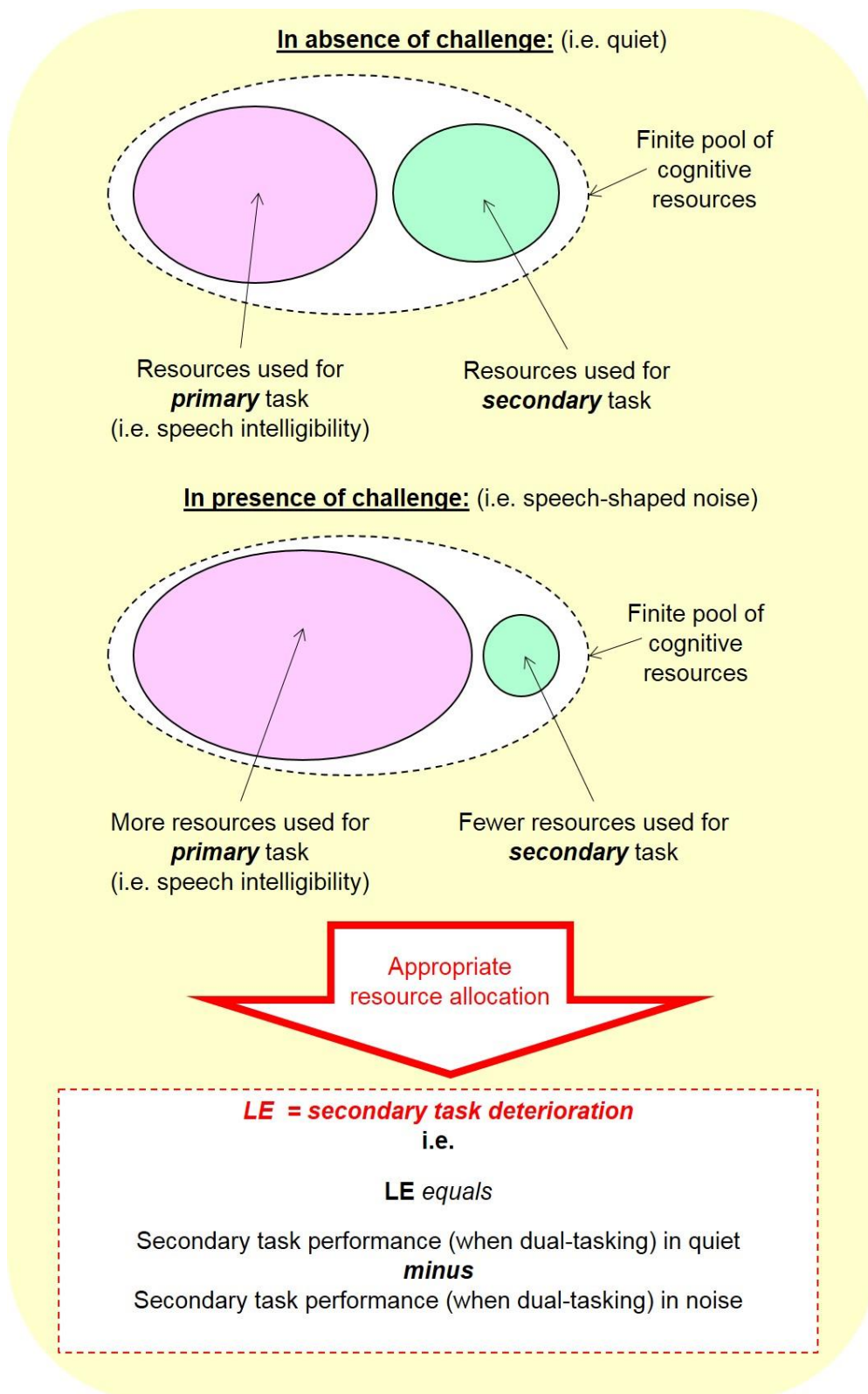


Figure 2.1: Two Venn diagrams showing the theoretical principles of how cognitive resources are allocated to the primary and secondary tasks within the dual-task paradigm in the absence of challenge (top Venn diagram) and presence of challenge (bottom Venn diagram), and how the LE calculation can be obtained from the change in resource allocation (red box).

2.3. Hypotheses

When acoustic challenge (i.e. speech-shaped noise) is present during the performance of the recall dual-task:

- Visual accuracy (of the secondary task) will deteriorate (i.e. positive LE scores will be attained);
- The pupil (of the eye) will dilate (i.e. there will be physiological corroboration of the behavioural data);
- Perceived effort will increase (i.e. there will be subjective corroboration of the behavioural data).

All of these changes will become greater when auditory stimuli become degraded (i.e. within the CI user group, or via the use of CI simulations in the normal hearing controls).

Finally, all of these changes will not be at the cost of primary task stability within the recall dual-task:

- Auditory accuracy will not change between the single-task or dual-task conditions conducted in noise.

2.4. Ethics and screening

All experiments in this thesis were run with the approval of the UCL Research Ethics Committee through its devolved procedure for innocuous experiments on adults (Code: SHaPS-2015-SR-016). Informed consent was obtained prior to the testing sessions, and a debriefing followed completion of all testing sessions. Furthermore, for all experiments, all normal hearing participants were screened for hearing loss exceeding 20dB (at octave frequencies between 250Hz to 8,000Hz) via pure-tone audiometry conducted in accordance to standardised clinical practice (MAICO audiometer, PC Werth).

2.5. Pilot Study

2.5.1. Methods

2.5.1.1. Participants

Four adult experienced CI users (4 females, mean age: 33.75 years, \pm 11.32 S.D) and six adult normal hearing (NH) controls (2 males and 4 females, mean age: 40 years, \pm 18.61 S.D) were recruited. All participants had normal, or corrected-to-normal vision, and were native British English speakers.

Hearing loss for all CI users was profound or total. See Table 2.1 and Table 2.2 for further information regarding the demographics of the CI users (such as aetiology of deafness and length of CI experience).

Table 2.1: Demographic detail of the CI users regarding: aetiology of deafness; type of deafness; length of experience with their CI; and the type of CI user.

CI user	Aetiology of deafness	Type of deafness	Duration of deafness (years)	Length of CI experience (years)	Type of CI user
1	Meningitis	Perilingual	22.5	21	Unilateral
2	Maternal rubella	Prelingual	24	7	Bimodal
3	Maternal rubella	Prelingual	42	10	Bimodal
4	Unknown	Prelingual	45	9	Bilateral

Table 2.2: Demographic detail of the CI users regarding: age, academic achievement level; language preference; and knowledge of sign language (either British Sign Language or Sign Supported English).

CI user	Age (years)	Academic achievement level	Language preference	Knowledge of sign language (BSL/SSE)
1	24	Postgraduate level	Speech	Yes
2	24	University level	Total Communication	Yes
3	42	Postgraduate level	Speech	Yes
4	45	University level	Speech	No

2.5.1.2. Procedure

Location of testing: All training and testing took place in a sound-proofed booth.

Primary task of the Recall Dual-Task: The primary task was a closed-set speech intelligibility task using sentences derived from the Children's Coordinate Response Measure (CCRM: Rosen, 2011; Bolia et al., 2000; Brungart, 2001). Sentences presented were in the following fixed format: "Show the [ANIMAL] where the [COLOUR] [NUMBER] is". When prompted, participants had to select the animal, colour and number (that they heard uttered) on a touchscreen (iPad Air 2 with WordPress Polldaddy plugin). There were six stimulus options available for each of the three words (animal: dog, cat, duck, pig, sheep, cow; colour: red, blue, green, pink, white, black; number: 1, 2, 3, 4, 5, 6).

Auditory stimuli for all participants (for both training and testing) were delivered free-field via a calibrated MS101II monitor speaker (Yamaha) such that the overall sound pressure level for auditory stimuli across all experimental conditions was 70dB SPL (including the conditions involving noise and/or vocoding).

To generate the CCRM sentences, a female speaker was recorded reading the sentences aloud. All auditory stimulus options (i.e. for animal, colour and number) had been chosen on the basis of being only one syllable in utterance, to ensure that no additional auditory cues were inadvertently provided to the participant. All sentences were equalised in duration (i.e. 2 seconds) by means of the PSOLA technique, as implemented by the 'respeed' function of the Speech Filing System software (Huckvale, 2013).

Secondary task of the Recall Dual-Task: The secondary task was a closed-set visual recall task, whereby participants had to recall the nature and order of a sequence of three shapes that appeared within three boxes. There were six stimulus options available for each of the three shapes to be recalled: square, circle, triangle, star, crescent and cross. The sequences of the three shapes were designed such that these shapes were never exactly the same within a given trial (i.e. just two shapes were the same, or all shapes were different). Responses for each shape were again selected using the same iPad touchscreen used for the auditory response.

All visual stimuli were presented using the Experiment Builder software of the eyetracker (that was being utilised for executing pupillometry) on a Benq computer screen (50cm x 70cm). All shapes and boxes that were part of the secondary task of visual recall were grey in colour with black outlines.

The choice of grey was to assist with the pupillometry (see section below about other LE indices for more information), by ensuring reasonable illuminance levels such that the resting pupil size of the eye was as close to neutral as possible (i.e. neither overly dilated nor overly constricted). 740 Lux was the recorded illumination level of these stimuli. A red fixation cross was provided within the central box to guide the participant's eye gaze.

Acoustic challenge: The mask of speech-shaped noise was utilised to manipulate the difficulty level for the primary task. To generate the speech-shaped noise, the long-term average speech spectrum was estimated from prior measurements for combined male and female voices (Byrne et al., 1994). The rms level per 1/3 octave band was converted into a spectrum level, from which a three-line approximation was utilised to characterise the speech-shaped noise (with the low-frequency portion rolling off below 120Hz at 17.5dB per octave, and the high frequency portion above 420Hz rolling off at 7.2dB per octave, with a constant spectrum portion in between). MATLAB (R2013b, MathWorks) was then utilised to add the speech-shaped noise to the CCRM sentences.

CI simulation training: Before testing of the Recall Dual-Task began, all NH controls underwent computer-based CI simulation training to be introduced to, and familiarised with, the vocoded sound. Software for this training was modified from that used by Faulkner et al. (2012). Training took 30 minutes in total: 15 minutes was in quiet; and 15 minutes in speech-shaped noise at the signal-to-noise ratio of 10dB. The stimuli consisted of a female speaker reading aloud extracts from "Indiana Jones and the Kingdom of the Crystal Skull". After each extract, 1-4 target words, along with the same number of foil words, were displayed on the computer screen. Participants were then required to select 1-2 words out of these options presented on the screen. Target words were always content words and the foil words shared at least two phonemes with the target.

If the participant selected a foil, the phrase was immediately replayed and the participant was required to choose again. This process continued until all the target words had been selected. After correct selection of all target words (be it on first or subsequent attempts), the extract was visually displayed on the computer screen and uttered once more, before proceeding onto the next section.

Vocoding parameters: For all test conditions involving CI simulations (i.e. the speech type of "vocoded") and also the CI simulation training, noise-excited vocoding was applied to the auditory stimuli using MATLAB (R2013b, MathWorks).

2.5. Pilot Study

The process of noise-excited vocoding preserved amplitude and temporal cues, varying primarily the amount of spectral information. The vocoding process used in the present study was modified from that used by Shannon et al. (1995). Six analysis filters, spanning 100-7500Hz and spaced at equal basilar membrane distances (according to Greenwood's (1990) cochlear position map) were used. An envelope was extracted from each analysis band using full-wave rectification and a 100Hz low-pass filter. Each band envelope was then multiplied by an independent white noise. The resulting modulated noises were passed through the 6 analysis filters and finally summed together to produce the vocoded auditory stimuli.

The choice of 6 filters was based on prior research intimating that 8 channels reflect the effective number of channels seen in CI users who perform relatively well in noise (Friesen et al., 2001). Therefore, in an attempt to capture the experience of the low performing CI users, 8 channels was reduced to 6.

Other LE indices:

Subjective Ratings: Upon completion of each test condition of the Recall Dual-Task, the participant's perceived effort was collected using subjective ratings in the form of a 5-point Likert scale. This scale had the following values: 1= Not at all hard work, 2= Quite hard work, 3= Medium hard work, 4= Very hard work, and 5= Extremely hard work. No ratings midway between the five available numbers were allowed. The subjective ratings were selected using the same iPad touchscreen used to collect the auditory and visual responses. For dual-task conditions, independent subjective ratings were obtained for the primary and secondary tasks.

Pupillometry: An Eyelink 1000 Plus eyetracker (SR Research) was used to record and estimate pupil size throughout all testing at a sampling rate of 500Hz. Due to pupil size changes being mediated by slow muscle fibres, a 500Hz sampling rate was sufficient for sensitive capture of pupil size fluctuations (Kramer et al., 2015).

The eyetracker system comprised a high speed video camera with a 35mm lens, which sent images of the eye to a host PC. The beta version of the Experiment Builder software (SR Research) on the host PC coordinated both the auditory and visual stimulus presentation with the pupil size recording, via a bespoke programming script. This script also analysed the images in real time, determining the size and location of the pupil and corneal reflection (i.e. 1st Purkinje image). However, this system could only provide highly sensitive estimates of pupil size provided that the head of the participant was stabilised on a chin rest, and also that a fixation cross was used.

2.5. Pilot Study

The use of the chin rest was in order to avoid pupil image distortion produced by foreshortening (caused by the eye rotating away from the camera: Hayes & Petrov, 2015). Thus, prior to executing each test condition in the Recall Dual-Task, all participants were set up on a chin rest 50cm away from the Desktop Mount, with the eye tracking camera focused on their eye.

As part of the attempt to optimise the pupil recording, an "autothreshold" procedure was also performed. This established the criteria that the eyetracker used in order to extract the parts of the recorded image associated with the pupil (versus those associated with corneal reflection). Pupil and corneal reflection thresholds comprised a greyscale value between 0 (i.e. black) and 255 (i.e. white). It was endeavoured that, at all times, the pupil and corneal reflection thresholds were within the ideal ranges of 75-125 and 200-240 respectively. Pupil size thus corresponded to the pupil area as calculated by the number of threshold pixels captured by the eyetracker's camera (i.e. these units were arbitrary and did not correspond to any physical units such as millimetres).

Calibration and validation were executed before each test condition of the Recall-Task. Calibration involved recording the position of the pupil and corneal reflection on the camera sensor at 9 target locations (i.e. a 3 x 3 grid) on the Benq computer screen that was being used for visual stimulus presentation. This allowed the eye tracking software to build a mathematical regression model to convert raw data (which was based in camera sensor pixel co-ordinates) into gaze data (which was based in screen pixel co-ordinates). A validation procedure presented the 9 targets for a second time to allow confirmation that the calibration model was sufficiently accurate (i.e. average error across the 9 points was always less than 0.5 degrees, and maximum error at any one point was less than 1 degree).

If validation was successful, a drift check was then executed, with the presentation of a single central target for the participant to fixate on. This was to ensure that there had been no accidental movement of the head in between setup and calibration and the beginning of a test condition (small changes in head position during calibration can translate into larger absolute inaccuracies in eyetracking). If this drift check was passed successfully, recording of pupil size (for the given test condition) could then begin. Only one eye was required in this recording, as changes in pupil size (and eye movements) are typically conjugate (Kramer et al., 2015).

Upon completion of each recording, any instances of blinks or saccades were removed. This is because pupil size measurements are no longer accurate during these occasions, either due to the obscuring of the pupil (i.e. during blinking), or the eye movements transgressing the calibrated limits (i.e. during a saccade: Hayes & Petrov, 2015). If these artefacts exceeded 15% of the dataset, this dataset was completely removed (as recommended by Kramer et al., 2015).

The effect of test condition on pupil size was derived by calculating the task-evoked pupil response. This required two types of pupil size recordings to be collected per trial: the “baseline” (i.e. when the pupil was in a resting state, with the eye focused on the fixation cross) and the “response” (i.e. the reaction of the pupil time-locked with the participant’s processing of the visual stimuli that had just been presented).

The peak value of pupil size within the “response” epoch was then identified, at which point a baseline correction was applied. This baseline correction entailed subtracting the median pupil size from the peak value. This median pupil size was derived from the “baseline” epoch of the trial. This baseline correction is needed because of the noisiness typically inherent to pupillometric data, produced by the natural micro-fluctuations in pupil size which can be present even at rest (Kramer et al., 2015).

Once all the trials’ baseline-corrected peak values were collected, the mean could then be computed to produce the task-evoked pupil response for the given test condition. This analysis technique is sometimes referred to as Peak Value Analysis, which is commonly chosen when the pupil response is being used to supplement other LE indices (such as subjective ratings, or other physiological measures: Kramer et al., 2015).

Test conditions (and trial structure): In total, there were 12 possible test conditions for stimulus presentation in the Recall Dual-Task for the NH controls (Figure 2.2), of which there were two conditions for speech type (i.e. normal, or vocoded as part of the CI simulations), three conditions for task type (i.e. visual task only, auditory task only, or dual-task), and two conditions for noise type (i.e. quiet or speech-shaped noise).

For the CI users, only 6 conditions were applicable, because they were not required to perform the test with CI simulations (i.e. the speech type of “vocoded”). For each test condition, 30 trials were presented. When speech-shaped noise was applicable, SNRs of -5dB and 0dB were used for the normal and vocoded conditions of speech type respectively.

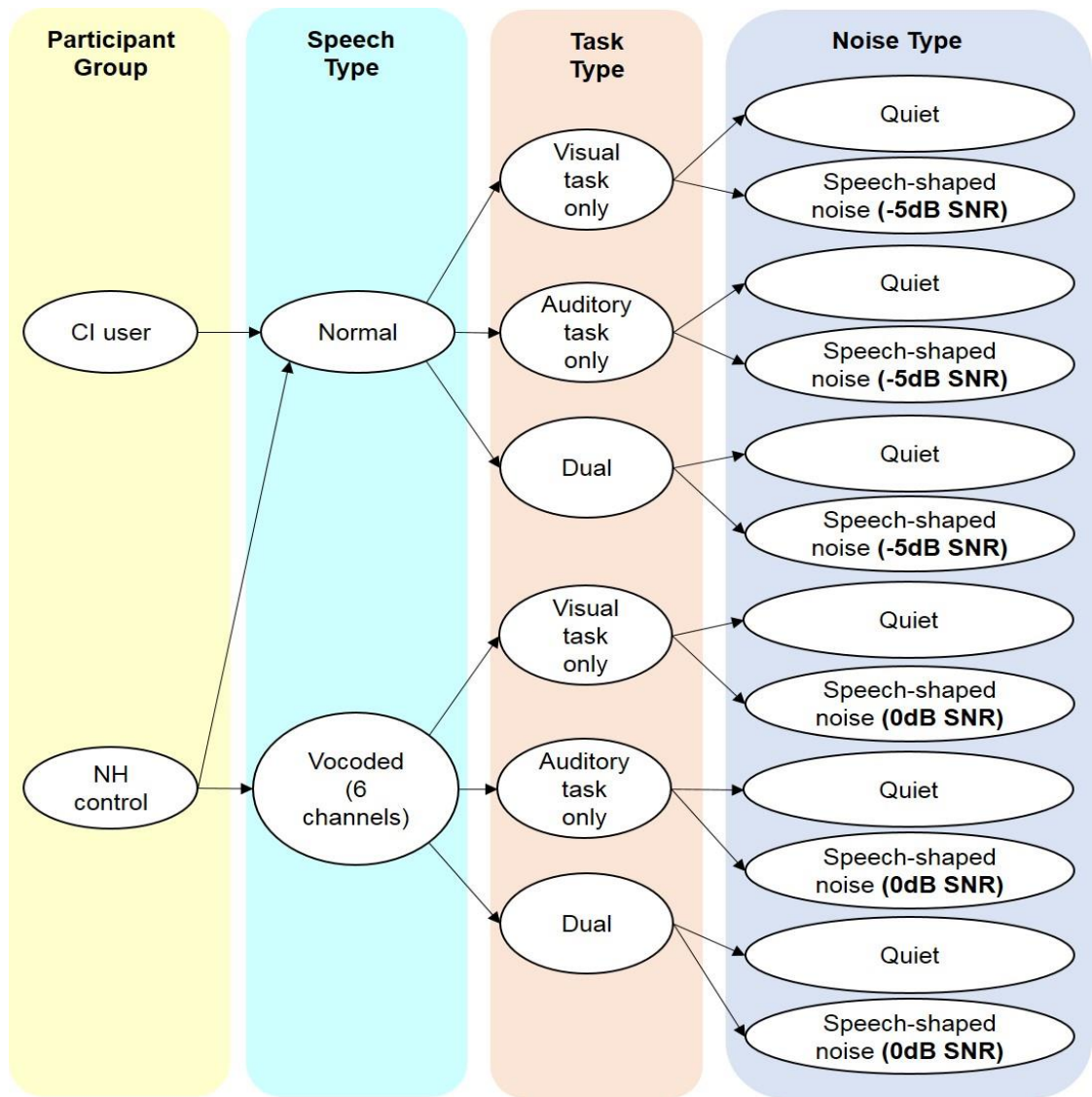


Figure 2.2: A schematic showing the breakdown of all possible twelve test conditions for stimulus presentation within the Recall Dual-Task, divided into participant group, speech type, task type, and noise type.

Each trial within a test condition had the following stages: a 5-second interval where instructions specific to the given experimental condition were provided regarding the nature of task type (i.e. whether it was visual-only, auditory-only or dual); a 3-second interval allowing for preparation of stimulus onset, with the participants' gaze directed to the fixation cross (during which the baseline measurement of the pupil was obtained); a 1-second presentation of visual stimuli; a 2-second presentation of auditory stimuli; a 3-second interval for retention of auditory and visual stimuli (during which any changes in pupil size in response to the processing of stimuli were collected); finishing with the response stage where the participants could provide their responses on the iPad touchscreen (Figure 2.3).

2.5. Pilot Study

In all test conditions, both auditory and visual stimuli were presented regardless of task type (i.e. visual-only, auditory-only, or dual). When participants were required to dual-task, the auditory response was collected before the visual response. When speech-shaped noise was applicable, this noise was presented on its own during the initial 3-second preparation, so that the participants' pupil size could acclimatise to the mask before the actual target stimuli were presented. The pupil exhibits an “alerting” response to the presence of any novel stimulus, whereby a transitory dilation (lasting 500 milliseconds) occurs (Kramer et al., 2015). It also needs to be noted that the duration of the retention interval (i.e. 3 seconds) is generally acknowledged to provide ample time for a pupil response to manifest (if one should occur: Kramer et al., 2015).

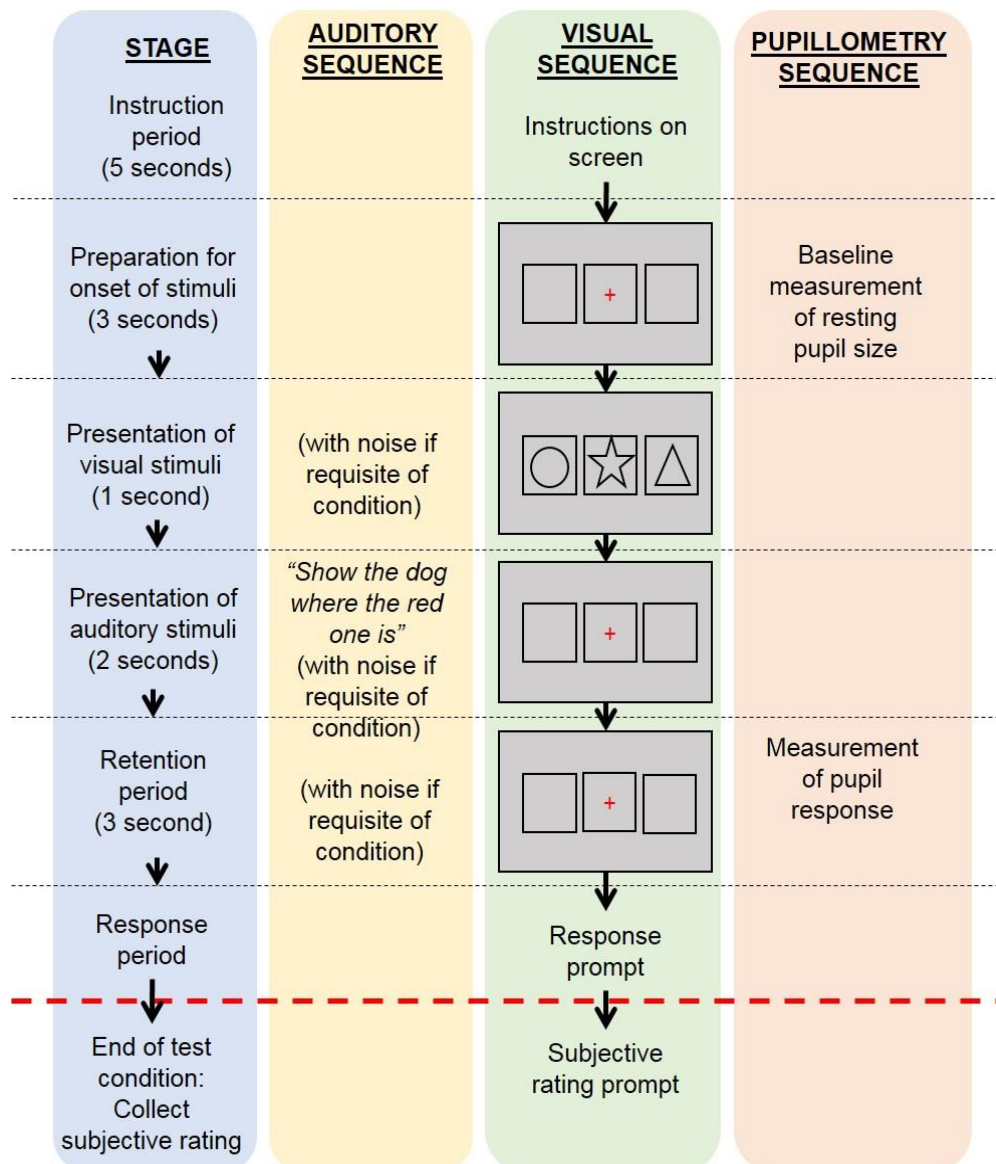


Figure 2.3: A schematic showing the structure of how each trial was presented during its thirty iterations within each test condition of the Recall Dual-Task. Detail is provided regarding the sequences of auditory and visual presentation, as well as the nature of the pupil recording and when subjective ratings are collected.

Additional assessment: To monitor whether there was reasonable neurocognitive similarity within and between participant group, all participants were screened in a separate testing session (for protocol of this assessment session, see Appendix 2.1) using the following tests/test batteries:

- Cambridge Neuropsychological Test Automated Battery: CANTAB (Cambridge Cognition)
- A modified version of Reading Span Test
- Weschler's Abbreviated Scale of Intelligence

The CANTAB was presented via the CANTAB Connect software on an iPad. The CANTAB is a well-validated battery that has been extensively used and standardised in a wide range of neurotypical and clinical populations (including CI users: Cambridge Cognition, 2006; Surowiecki et al., 2002). A particular advantage of the CANTAB is that a relatively broad and useful assay of cognitive and executive function can be executed in a reasonable period of time (i.e. an hour, including training) in a format that is equally accessible to both CI users and NH controls (i.e. visually).

The CANTAB comprises five vision-based sub-tests in the following order: the Motor Screening Task, Reaction Time Index, Paired Associates Learning, Spatial Working Memory, Rapid Visual Information Processing, and Delayed Matching to Sample. These tests collectively enabled the assessment of sensorimotor function, motor speed, mental speed, impulsivity, episodic visual memory, spatial working memory, cognitive strategizing, sustained attention, perceptual matching, short-term visual memory, as well as the ability to learn (for more detail about what each of these sub-tests entailed, see Appendix 2.2). Immediately before completion of these sub-tests, all participants underwent training (which took approximately 30 minutes), whereby they were taken through each sub-test and an explanation was given as to what was expected of them. Any oral instructions were provided in written form for CI users within a transcript. Only data from the testing phase were utilised.

The Reading Span test is designed to tax working memory storage and processing simultaneously, thus giving an index of working memory capacity, as well as providing a check of any severe linguistic problems (Daneman & Carpenter, 1980; Baddeley et al., 1985). A modified form of the English version of Ronnberg's Reading Span task was utilised to provide a computerised version of this Reading Span test (Ronnberg, 1989; Schoof & Rosen, 2014).

2.5. Pilot Study

The participant's task in the Reading Span was to read a sequence of sentences and then recall either the first, or the final, word of each sentence, ideally in correct serial order. The sequence of sentences was presented word by word, at a rate of one word per 0.8s. Half of the sentences were absurd, such as "The train sang a song", and half were sensible sentences, such as "The girl brushed her teeth". The participant was not informed as to whether it was the first, or the last, word being requested in the recall until the presentation of a given sequence of sentences had been completed.

Throughout the course of the testing, the sequence increased incrementally in length from three sentences to six sentences. There were two types of scoring to determine the participants' working memory capacity: the percentage score of correctly recalled words in the correct order; and the percentage score irrespective of order.

The WASI (Pearson) was utilised as a measure of intellectual function (Wechsler, 1999). The WASI has four subtests available, two of which are the verbal tests (i.e. measures of crystallised abilities, such as word knowledge, verbal concept formation, fund of knowledge and verbal reasoning) and performance tests (i.e. measures of non-verbal abilities). Only the performance sub-tests (i.e. the Block Design and the Matrix Reasoning) were performed in this study, to ensure equal access for CI users and NH controls. Participant scores were converted into standardised t-scores and summated to derive the non-verbal intelligence quotient (IQ). This non-verbal IQ is designed to encompass the ability to analyse and synthesise abstract visual stimuli, abstract reasoning, non-verbal concept formation, visual perception, simultaneous visual processing, figure-group separation, as well as perceptual learning, visuomotor coordination and cognitive organisation.

These neurocognitive assessments (i.e. CANTAB, Reading Span and WASI) were intended to act as a screening procedure in order to identify any "supertaskers" (with superior multitasking skills) who would be unrepresentative of the general population (Watson & Strayer, 2010). However, none were revealed within the recruited sample for the Pilot Study; instead, it was found that reasonable neurocognitive similarity was achieved both within and between the two participant groups (for data, see Appendices 2.3 and 2.4).

In addition to this neurocognitive assessment, a modified version of the Children's Coordinate Response Measure (CCRM) was also performed to confirm speech reception thresholds (SRTs) (Rosen, 2011; Bolia et al., 2000; Brungart, 2001). The CCRM is an adaptive tracking procedure (which utilised the same sentence corpus as the Recall Dual-Task) that is designed to find the SNR at which the participant achieves 50% auditory accuracy.

2.5. Pilot Study

The CCRM was performed three times in normal listening conditions, and also three times in vocoded listening conditions (for NH controls only). Where vocoded stimuli were applicable, it was ensured that noise-excited vocoding parameters matched those used in the CI simulation training and the Recall Dual-Task. An average was taken of the three SRTs. The starting SNR for both versions of the CCRM was 20dB, with initial decrement in SNR being 9dB before the tracking procedure was activated. Incorrect responses were initially ignored for the first three trials (this occurred for both the normal and vocoded versions of the CCRM). However, should the participant persist in giving incorrect responses at the maximum SNR, the test stopped and the participant was reported to have failed that particular attempt of the CCRM. The maximum number of trials allowed for the entire CCRM was 25.

The order in which all these assessments (i.e. the CANTAB, Reading Span, WASI and CCRM) were executed were counter-balanced across participants, with sub-tests also being counter-balanced where possible.

Protocol: Before the test conditions of the Recall Dual-Task were implemented, all participants underwent a practice run (this practice run was after the CI simulation training for the NH controls). This practice run involved executing 9 trials of the following conditions: visual task only (in both quiet and noise), auditory task only (in both quiet and noise), and dual-task (in both quiet and noise). Within this practice run, participants were also exposed to all possible stimulus options for both visual and auditory stimuli.

For the overall protocol of the testing session for the Recall Dual-Task, see Table 2.3. To control for fatigue and/or training effects, the order of test conditions (i.e. steps 4-15 in Table 2.3) were randomised and counter-balanced. Furthermore, the sequences of the auditory and visual stimuli themselves were randomised, such that there was equal frequency of each stimulus option and no repetition of auditory or visual stimuli (as well as no repetition in the pairings of auditory and visual stimuli) both within and across all test conditions.

Statistical analysis: Due to the small sample size, only descriptive statistics were pursued.

2.5. Pilot Study

Table 2.3: Protocol for the testing of the Recall Dual-Task. Detail is provided regarding whether it is training, a practice run or a test run, as well as the task type, test condition, speech type (i.e. normal or vocoded), noise type (i.e. quiet or in speech-shaped noise), and also number of trials presented (or duration of time if applicable). Note that CI users did not undergo CI simulation training, or any conditions involving vocoded stimuli. Red numbers with asterisks indicate all the steps in the protocol where the order of the test conditions were randomised and counter-balanced together. Pink shading indicates where training was implemented. Blue shading indicates application of vocoding.

Step	Training, Practice or Test?	Task Type or Test Condition	Speech Type	Noise Type	Total number of trials/ Total duration
1	Training	Indiana Jones	Vocoded (6 channels)	Quiet	15 minutes
2	Training	Indiana Jones	Vocoded (6 channels)	Speech-shaped noise at 10dB SNR	15 minutes
3	Practice	Visual task only	N/A	Quiet	9 trials
	Practice	Auditory task only	Normal	Quiet	9 trials
	Practice	Dual	Normal	Quiet	9 trials
	Practice	Visual task only	N/A	Speech-shaped noise at -5dB SNR	9 trials
	Practice	Auditory task only	Normal	Speech-shaped noise at -5dB SNR	9 trials
	Practice	Dual	Normal	Speech-shaped noise at -5dB SNR	9 trials
4 *	Test	Visual task only	N/A	Quiet	30 trials
5 *	Test	Auditory task only	Normal	Quiet	30 trials
6 *	Test	Dual	Normal	Quiet	30 trials
7 *	Test	Visual task only	Normal	Speech-shaped noise at -5dB SNR	30 trials
8 *	Test	Auditory task only	Normal	Speech-shaped noise at -5dB SNR	30 trials
9 *	Test	Dual	Normal	Speech-shaped noise at -5dB SNR	30 trials
10 *	Test	Visual task only	N/A	Quiet	30 trials
11 *	Test	Auditory task only	Vocoded (6 channels)	Quiet	30 trials
12 *	Test	Dual	Vocoded (6 channels)	Quiet	30 trials
13 *	Test	Visual task only	N/A	Speech-shaped noise at 0dB SNR	30 trials
14 *	Test	Auditory task only	Vocoded (6 channels)	Speech-shaped noise at 0dB SNR	30 trials
15 *	Test	Dual	Vocoded (6 channels)	Speech-shaped noise at 0dB SNR	30 trials

2.5.2. Results and Discussion: Descriptive statistics

2.5.2.1. *LE measurement (secondary task performance): Visual accuracy*

When visual recall across all test conditions was examined (Figure 2.4 a, b and c), there was a tendency for accuracy to decrease when dual-tasking was required. This trend was fairly consistent across participant groups, and also across listening conditions for the NH controls (i.e. normal versus vocoded).

When the LE scores (i.e. difference in percentage) were calculated for each participant (Figure 2.4 d), it was found that the majority of NH controls attained positive scores in both normal and vocoded listening conditions (LE = visual recall accuracy when dual-tasking in quiet minus visual recall accuracy when dual-tasking in noise).

In contrast, however, all CI users obtained negative LE scores (i.e. their visual accuracy improved when dual-tasking in noise), thus implying a decrease in their LE when speech-shaped noise was introduced. This is contrary to the original hypothesis, which stipulated that the degraded nature of the CI input would engender comparatively higher LE levels to that experienced by the NH controls. Whilst it is possible that the LE measurement achieved by the recall dual-task is valid, and the decrease in LE genuine, the other indices of LE (that had been concurrently recorded) warranted closer scrutiny.

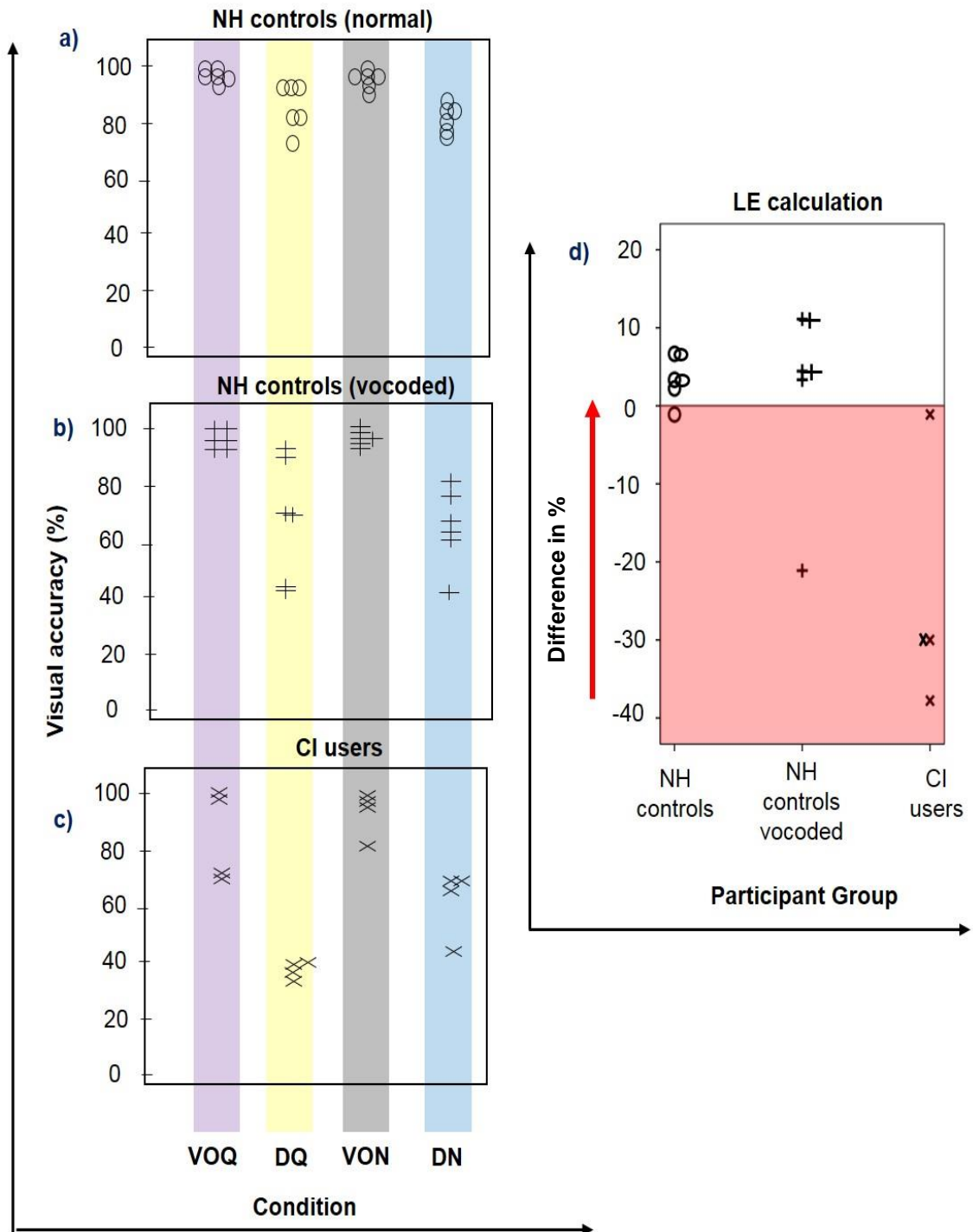


Figure 2.4: Dot plots of visual recall accuracy (**a, b, c**) and LE (**d**). VOQ = Visual task only in quiet, DQ = Dual-tasking in quiet, VON = Visual task only in noise, DN = Dual-tasking in noise. Red arrow indicates direction in which LE increases. Red shading indicates when LE has decreased in presence of noise.

2.5. Pilot Study

2.5.2.2. Evaluation of validity of LE measurement: Pupil response

Due to difficulties in calibration, the entire pupillometry dataset had to be excluded for two NH controls. Therefore, for the purpose of descriptive analysis of pupil size, the datasets of 4 NH controls and 4 CI users were utilised.

In general, there was considerable variability in the mean task-evoked pupil responses across participants in both groups (Figure 2.5). However, the pupil size for the majority of participants (both NH controls and CI users) did appear to increase (relative to resting pupil size) when expected to perform a single task, i.e. the visual-only and auditory-only conditions (for both quiet and in noise). For the NH controls, in both the normal and vocoded listening conditions, the pupil size then tended to increase further still when dual-tasking was required (both in quiet and in noise). This data trend was consistent with the original hypothesis that pupil size increases as LE increases with the introduction of an acoustic challenge.

For CI users, however, dual-tasking did not seem to elicit the same enlargement in pupil size; instead, pupil responses appeared diminished (with pupil constriction even being observed in some cases), particularly when dual-tasking in noise. Thus, once again, the response of the CI users opposes that hypothesised.

There remained one more branch of the across-methods triangulation attempted in this study: the subjective ratings. Perusal of the subjective ratings is particularly important because of a caveat with the pupil response data: not only was there substantial variability in pupil size between participants, there was also inconsistency found in pupil size within participants (across the trials of each test condition). This inconsistency applied to the pupil size at rest, as well as during the task-evoked responses (see Appendix 2.5 and 2.6 respectively).

While it is likely that this variability is due to the intrinsic (and natural) fluctuations typically found in pupil size, caution is still needed in data interpretation (Kramer et al., 2015; Zekveld & Kramer, 2014).

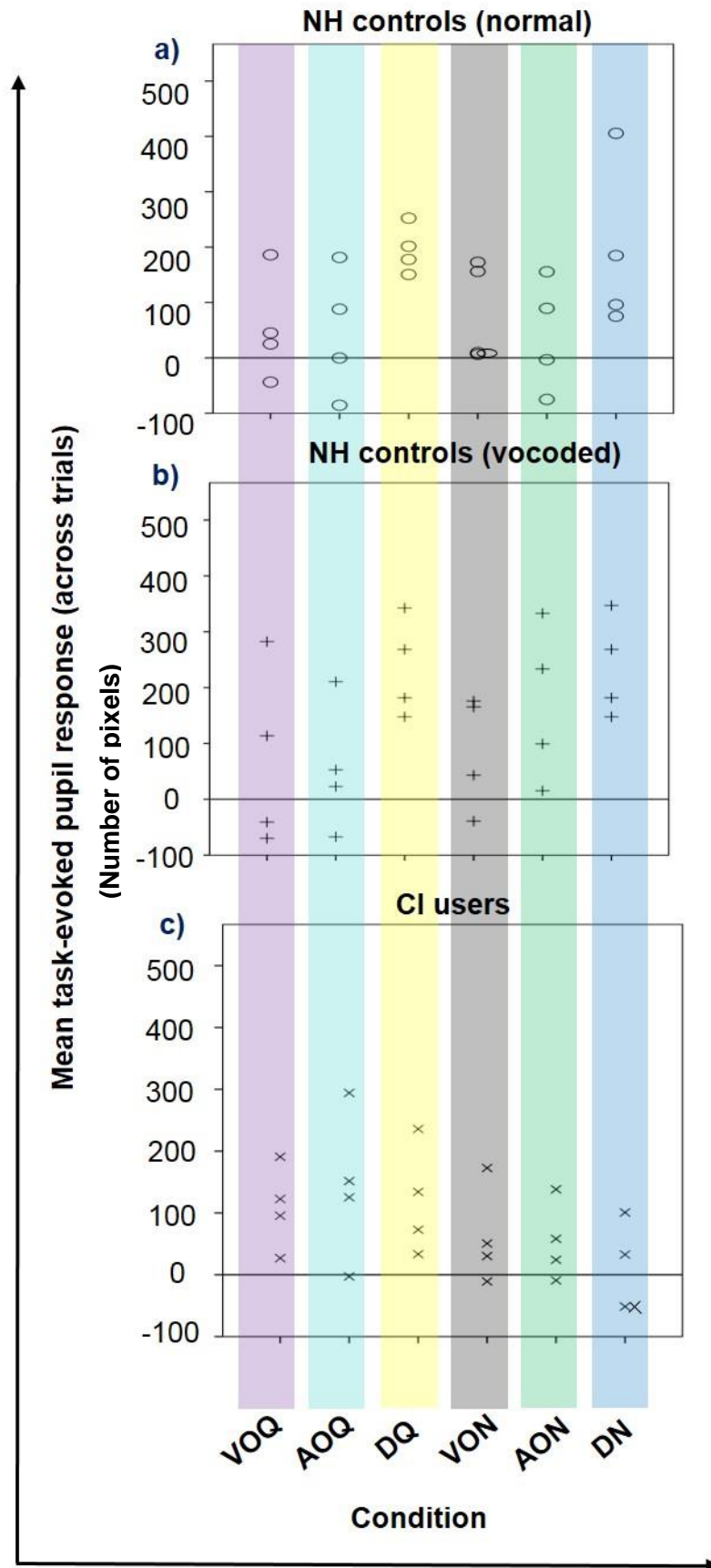


Figure 2.5: Dot plots of mean task-evoked pupil response. VOQ = Visual task only in quiet, AOQ = Auditory task only in quiet, DQ = Dual-tasking in quiet, VON = Visual task only in noise, AON = Auditory task only in noise, DN = Dual-tasking in noise.

2.5.2.3. *Evaluation of validity of LE measurement: Subjective ratings*

Overall, there appeared to be a trend that, as noise or vocoding (or both) were introduced, subjective ratings shifted toward “extremely hard work” (Figure 2.6). This trend was true whether the participant (NH control or CI user) was performing the visual or the auditory tasks on their own, or dual-tasking.

However, the comparison that was of particular interest in these subjective ratings was that between dual-tasking in quiet versus dual-tasking in noise in the CI user group, particularly for the secondary visual task (i.e. DQV versus DNV in Figure 2.6). It was revealed that, whilst subjective ratings were always high within the dual-tasking conditions, there was an upward shift in rating: the entire CI user group claimed dual-tasking in noise to be “extremely hard work”. This is precisely as hypothesised. There is thus little consistency between the behavioural, physiological and subjective data for the CI users.

Several explanations can potentially be offered to account for this lack of corroboration. One possibility is that the subjective ratings were describing a different aspect of LE to that detected by the recall dual-task, or indexed by pupil size. Alternatively, it is feasible that the subjective ratings were an inconsistent measurement of LE. Indeed, there was variability in the subjective ratings given by all participants (both NH controls and CI users) across all test conditions. This variability may arise from inter-individual variability and/or intra-individual variability in how the participants subjectively perceive listening effort.

Another factor to consider is that the ratings themselves are subject to interpretation. When asking participants to gauge their perceived effort, they may either interpret this as to how much effort they actually used, or they may give an answer that relates to how much effort they believe they would need to apply (in order to succeed in performance of the given test condition).

Yet another aspect to consider is that the subjective ratings may have been unreliable. This concern arises from the anecdotal reports (during debriefing) of participants finding the requirement to give their subjective ratings for the primary and secondary tasks separately within the dual-task conditions confusing.

However, there is another conceivable explanation which is more concerning: resource allocation within the recall dual-task had become inappropriate. It may be the case that the difficulty level caused by speech-shaped noise became excessive (as indicated by the high ratings of perceived effort), which then encouraged the CI users to disengage with the primary task. This would, in turn, lead to the liberation of cognitive resources for the secondary task, hence the bolstered visual accuracy scores.

The diminished pupil response is also consistent with this notion of task disengagement: pupil constriction has previously been associated with cognitive overload, whereby the breach of cognitive capacity caused the abandonment of the task at hand. The subsequent removal of task demand then ameliorated the cognitive duress that would have triggered pupil dilation, leading instead to pupil constriction (Zekveld & Kramer, 2014; Hopstaken et al., 2015).

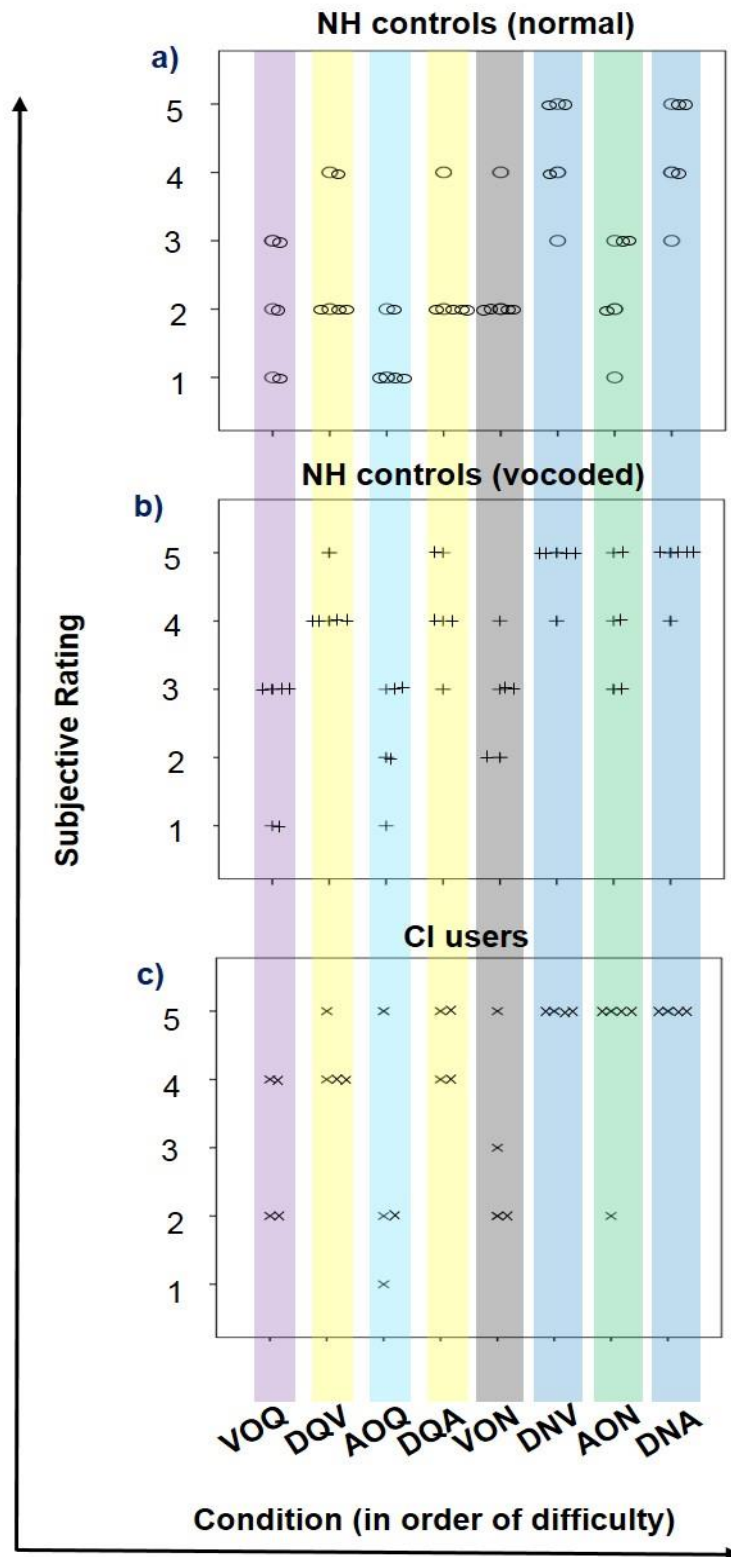


Figure 2.6: Dot plots of subjective ratings. VOQ = Visual task only in quiet, DQV = Visual task when dual-tasking in quiet, AOQ = Auditory task only in quiet, DQA = Auditory task when dual-tasking in quiet, VON = Visual task only in noise, DNV = Visual task when dual-tasking in noise, AON = Auditory only in noise, DNA = Auditory task when dual-tasking in noise. Subjective ratings: 1= Not at all hard work, 2 = Quite hard work, 3 = Medium hard work, 4 = Very hard work, 5 = Extremely hard work.

2.5.2.4. *Stability of primary task performance: Auditory accuracy*

For all NH controls, there was the tendency for auditory accuracy to decrease when speech-shaped noise was present (Figure 2.7 a, b and c), although this was minor within the normal listening condition (as indicated by performance levels remaining close to ceiling). CI users, on the other hand, exhibited more marked deterioration in noise. However, there was substantial variability in the CI users' performance, in quiet as well as in noise.

To better ascertain the likelihood of inappropriate resource allocation having occurred, primary task stability was calculated for each participant (by subtracting auditory accuracy when performing only the auditory task in noise from auditory accuracy when dual-tasking in noise – see Figure 2.7 d). It was discovered that some of CI users demonstrated some deterioration in auditory accuracy when dual-tasking in noise.

Furthermore, the majority of NH controls also exhibited deterioration in auditory accuracy when dual-tasking in noise, irrespective of listening condition (i.e. normal or vocoded). In addition, when the LE scores (as calculated from the visual accuracy scores when dual-tasking) were plotted as a function of primary task stability (Figure 2.7 e), three of the participants whose performance had deteriorated (when dual-tasking in noise) also attained negative LE scores.

Whilst the negative data points (in both auditory accuracy and LE) may simply be due to measurement variability, there is the need to proceed with caution because it is not possible to discount the likelihood that the participants' resource allocation may have become awry during dual-tasking (with NH controls being just as culpable of this as the CI users).

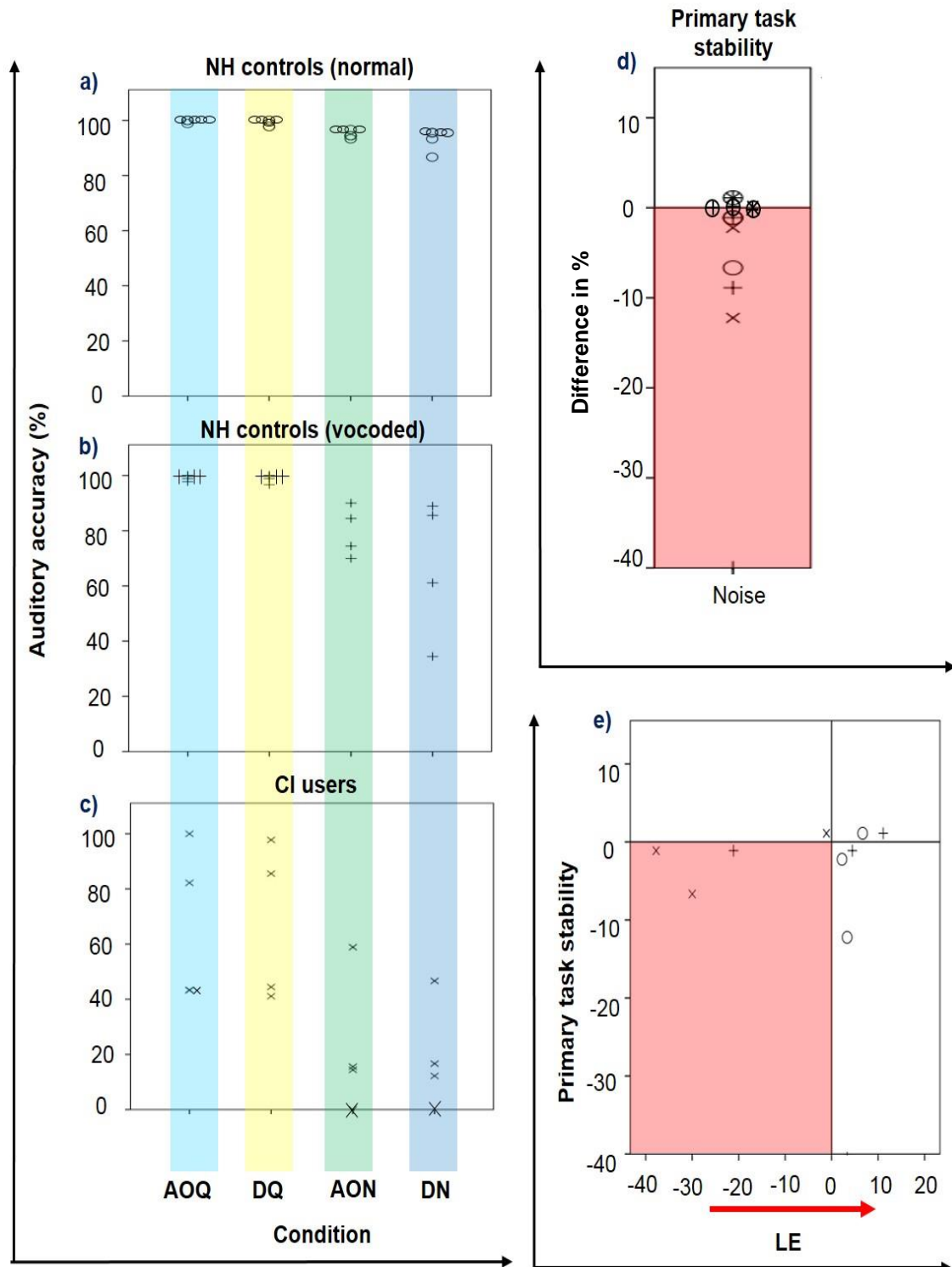


Figure 2.7: Dot plots of auditory accuracy (**a, b, c**) and primary task stability (**d**), and scatterplot of LE versus primary task stability (**e**). AOQ = Auditory task only in quiet, DQ = Dual-tasking in quiet, AON = Auditory task only in noise, DN = Dual-tasking in noise O = NH control (in normal listening condition), + = NH control (in vocoded listening condition), X = CI user. Red arrow indicates direction in which LE increases. Red shading indicates when variable has decreased in noise.

Units of primary task stability & LE = Difference in percentage.

2.5.2.5. Evaluation of viability of LE measurement

As part of the attempt to elucidate what might have triggered the disengagement from the primary task, the participant's speech-in-noise data was revisited (Figure 2.8). An adaptive version of the CCRM (i.e. using the same corpus of auditory stimuli as implemented within the recall dual-task) had been performed as part of the neurocognitive profiling, from which speech reception thresholds could be calculated.

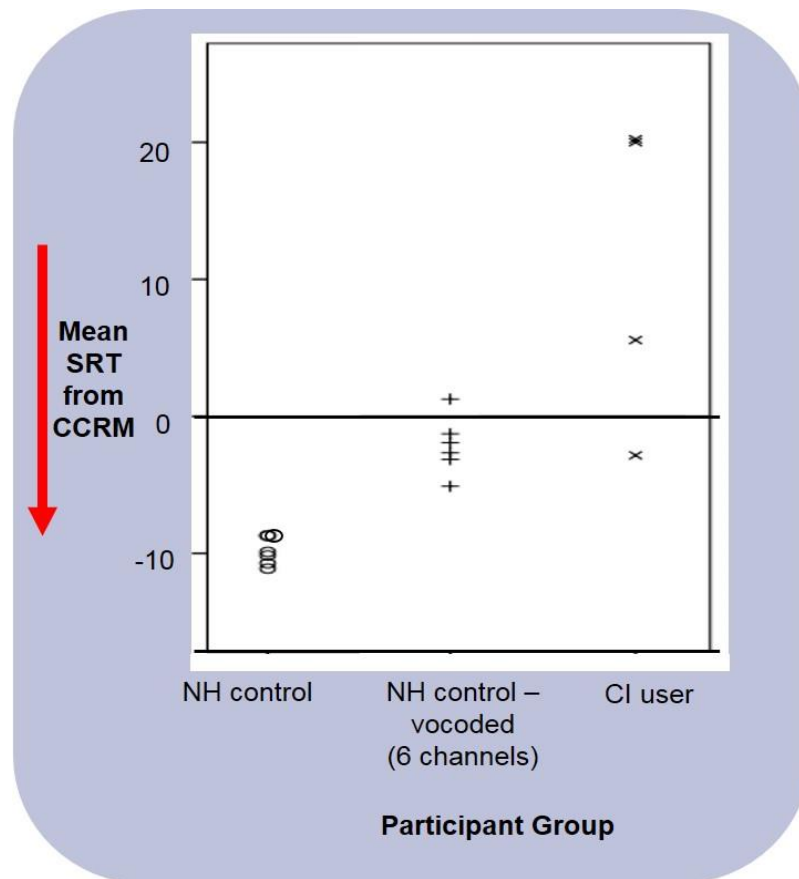


Figure 2.8: Dot plot of the mean speech-reception threshold (SRT) obtained from the three attempts of the CCRM for NH controls (in normal and vocoded listening conditions) and CI users. Red arrow indicates direction of superiority in SRT. Unit of SRT = dB.

It appeared that the SNR of -5dB might have been inappropriate for the CI users. This is because the speech reception thresholds achieved were greater than this SNR.

This occurred in spite of the fact that this SNR had been carefully chosen following consultation with clinical audiologists working with CI users, who believed that this level of challenge was both feasible and appropriate.

There is thus a tangible possibility that the participants, especially the CI users, did indeed become overwhelmed by the acoustic challenge presented by the noise. This may, in turn, account for those instances where apparent primary task instability coincided with a negative LE score.

However, it does need to be noted that the chosen noise level is not the only potential culprit for primary task instability. The secondary task of visual recall may also be causing an issue by being relatively more novel, more salient, or more difficult than the primary task, thus encouraging (and not necessarily consciously) the participant to focus instead on visual recall. The timing of the visual stimulus presentation may also have inadvertently contributed to rendering the secondary task more salient. This is because the visual stimulus occurred prior to the onset of the spoken sentence and may have diverted the participants' attention. This ordering in stimulus presentation was unavoidable, because it was necessitated by the requirements of the pupillometry measurement.

Regardless of the cause, if the secondary visual task has attracted more attention away from the primary auditory task, this would then lead to the same consequence of bolstered visual accuracy (and negative LE scores) when noise is present.

2.6. Conclusions of Pilot Study

Overall, the NH controls exhibited deterioration in secondary task performance (with reduced visual recall accuracy) when speech-shaped noise was introduced in both normal and vocoded listening conditions, suggesting an increase in LE. This was corroborated with the tendency for increased pupil responses and higher subjective ratings of perceived effort in noise too.

On the other hand, the CI users elicited negative LE scores, along with concomitant pupil constriction. This could reflect a genuine decrease in LE when noise was presented.

However, with the CI users' subjective ratings increasing in noise, as well as possible instability in primary task performance (in terms of their auditory accuracy), there is now a likelihood that inappropriate resource allocation may have occurred during the recall dual-task, with the secondary task being prioritised (instead of the primary task as instructed). This possibility of primary task instability in noise also applied to the NH controls (irrespective of whether it was normal or vocoded stimuli).

As a result, the viability of this recall dual-task as a robust LE measurement for CI users is now questionable.

The next chapter will explore the potential perpetrators responsible for this risk of aberrant resource allocation, and consider how this may be resolved by changes in the dual-task paradigm. Experiments with revised dual-task methodology will thus be undertaken in Chapter 3.

PART THREE:

**DIGIT STREAM
DUAL-TASK**

(Normal Hearing
Participants)

CHAPTER THREE:

Refining and testing a digit stream dual-task (Experiments 1, 2 & 3)

3.1. Introduction

As concluded in the previous chapter, the recall dual-task (that had been created and trialled in the Pilot Study) is of questionable viability as a potential test for LE. This is due to the tangible possibility that inappropriate allocation (of cognitive resources) had occurred during the execution of the dual-task paradigm. This then compromised the underlying theoretical principles that enable the use of secondary task performance to index LE. Potential origins of this incorrect allocation policy include both the primary and secondary tasks of the recall dual-task: either the noise levels had rendered the primary task excessively difficult, or the secondary task had inadvertently attracted more cognitive capacity than should have been allocated.

3.2. Proposed solution

Since the two potential causes of the incorrect resource allocation were not necessarily mutually exclusive (indeed, they might even be interacting to compound matters), both the primary and secondary tasks will be replaced with new tasks within another novel dual-task design.

In the case of the primary task, sentence-based stimuli are still pursued because of evidence suggesting that the use of sentences renders LE measurement more sensitive than single words (Gatehouse & Gordon, 1990). However, an open-set sentence intelligibility task will be utilised this time, in order to attempt to introduce more granularity into the auditory score.

In the case of the secondary task, the visual modality is still utilised in order to avoid any conflict in processing required by the primary and secondary tasks (Downs, 1982; Pals et al., 2013). However, rapid serial visual presentation of numerical digits will be applied. Rapid serial visual presentation is frequently used to examine short-term memory as well as the temporal aspects of attention (Coltheart, 1999; Intraub, 1980; Spence, 2002; Rees et al., 1999; Weichselgartner & Sperling, 1987). This methodology requires the participant to detect targets which are embedded within a continuous stream of visual items, which all appear (and disappear) at the same location (Potter, 1976, 1993).

3.2. *Proposed solution*

It has been postulated that this rapid serial visual presentation is ecologically valid, because humans are continually browsing their external environment (in both auditory and visual aspects) as part of ongoing monitoring, with rapid decisions constantly made about which details to retain for further processing (Potter, 1976, 1993; Spence, 2002; Spence & Witkowski, 2013). This methodology also offers the additional advantage of being more flexible in terms of the adjustment of the difficulty level; indeed, the speed of serial visual presentation can be manipulated with a high level of control. This, in turn, offers promise in terms of being able to optimise the sensitivity of this secondary task in LE measurement.

However, it does need to be noted that the use of the digit stream as the secondary task now contraindicates the use of pupillometry, because of the rapidly flashing stimulus. However, subjective ratings will still be pursued in order to help with the evaluation of the viability of this new dual-task design as a form of LE measurement.

As with the recall dual-task in Pilot Study, there is no precedent for this newly developed digit stream dual-task. Thus, there is a need to collect baseline data, especially for the new visual task. Therefore, two preliminary small-scale experiments (i.e. Experiments 1 and 2) were conducted in order to troubleshoot and refine the new test parameters, before progressing onto a larger-scale study to trial the new dual-task design with a cohort of 28 normal hearing participants (i.e. Experiment 3).

Because of the exploratory nature of Experiments 1 and 2, there will be no calculation of LE scores. Instead, there will be an attempt to ascertain the best speed for the rapid serial visual presentation. Only Experiment 3 will attempt to quantify LE. In addition, only normal hearing participants will be tested in these three experiments. This is in order to avoid any ambiguity of data interpretation that might be imbued by CI heterogeneity, such as differences in clinical characteristics (Blamey et al., 2013; Green et al., 2007; Holden et al., 2013).

3.3. Data collection

In order to enable recruitment of a wide variety of participants, as well as larger cohorts (in the attempt to more closely represent the general population), data collection for all experiments from this point onwards became peripatetic. Therefore, testing sessions were conducted within the participants' home and work environments.

To ensure that the testing conditions were still optimal, quiet rooms (with no distraction of other individuals, or activities) were always used. In addition, ambient (A-weighted) sound levels were continually monitored throughout data collection via the use of a sound level meter (ST-805 sound level meter, Reed Instruments: accuracy $\pm 1.5\text{dB}$).

Throughout Experiments 1-8, ambient sound levels never exceeded 38dBA (please note that a comparison recording in a sound-proofed testing booth, using the same equipment, obtained 35dBA).

3.4. Experiment 1

3.4.1. Hypotheses

When the participant is dual-tasking, visual accuracy (determined now by target detection percentage) will:

- Decrease relative to performing just the visual task on its own;
- Decrease further still when the SNR worsens.

In addition, these changes will not be at the cost of primary task stability:

- Auditory accuracy will remain constant between the single-task and dual-task conditions conducted in noise.

3.4.2. Methods

3.4.2.1. Participants

Six adult normal hearing (NH) controls (4 females and 2 males, mean age: 43.7 years, ± 16 S.D) were recruited. All participants had normal, or corrected-to-normal vision, and were native British English speakers.

3.4.2.2. Procedure

Location of testing: Testing sessions were conducted within the participants' homes.

Primary task of Digit Stream Dual-Task: The primary task was an open-set speech intelligibility task using sentences derived from the Basic English Lexicon (BEL: Calandruccio & Smiljanic, 2012). This sentence corpus comprised a total of 20 lists, each possessing 25 sentences.

These lists were equalised with regards to rate of occurrence of words in the lexicon, total number of affricates and fricatives, number of syllables, and distribution of syntactic structure. Prior testing with NH listeners indicated similar speech intelligibility across lists (Calandruccio & Smiljanic, 2012). However, this study used American talkers and listeners.

For the purposes of the digit stream dual-task, the sentence stimuli were re-recorded using a British native male speaker of SSBE (Standard Southern British English) reading the sentences aloud. The longest sentence was 2.55 seconds in duration (i.e. “Her grandparents are serious and sometimes cruel”), whilst the shortest sentence was 1.27 seconds (i.e. “The man ate fried chicken”). These sentences were presented via a Graphical User Interface (GUI) on MATLAB (R2015b, MathWorks).

Each sentence has 4 key words. When prompted, the participant was required to type what they heard within a response box on the GUI. Once they submitted the answer, the four key words were displayed within the GUI (providing immediate feedback). At this point, the participant was able to score their own accuracy by using a mouse cursor to click on the key words they had successfully typed.

The criterion for a correct key word was that the root of the word must have been successfully heard (and typed). For instance, if the key word “juiciest” was presented, the root of this was “juice”. Therefore, any derivatives based on this root was allowed, such as “juicy”, “juices” and “juicier”. However, should the participant’s typed response be similar to the target key word but entail a different vowel (e.g. the past participle of an irregular verb, such as “took” versus “take), the participant was not allowed to score themselves as correct. All responses were recorded, so that the participants’ marking could be checked afterwards. The percentage of key words detected across all trials was used to determine auditory accuracy for each test condition.

All auditory stimuli were delivered via calibrated headphones (HD380 Pro Collapsible headphones, Sennheiser) such that overall sound pressure level for auditory stimuli across all experimental conditions was 70dB (including the conditions involving noise). These headphones also produced a passive attenuation of 32dB, thus further assisting with the control of ambient sound levels.

Secondary task of the Digit Stream Dual Task: The secondary task entailed a rapid visual serial presentation of digits (from 0-9) forming a digit “stream”. The digits appeared (in black) and disappeared from the same location within a small white box set on a neutral grey background. A fixation cross was used at the beginning of each digit stream to direct the participant’s gaze.

All visual stimuli were also presented via a GUI (as generated by MATLAB), on a 15.6" HD widescreen of a Latitude 5440 laptop (Dell).

This digit stream was presented at different speeds, with the speed being determined by manipulating the total number of digits presented in the stream, which was of a fixed duration (i.e. 3.2 seconds). The participant was required to monitor the digit stream and count the incidence frequency of the target digit, which was the digit “3”. When prompted, the participant typed the number of targets they detected within a response box (provided by the same GUI presenting the visual stimuli). The percentage of targets detected across all trials was used to determine visual accuracy for each test condition.

In the event of the participant reporting more targets than actually were presented (within a given trial), the visual accuracy score would become skewed. Therefore, if this over-reporting occurred, these responses were corrected back to the actual number of targets presented. This was because these extra targets were assumed to be misidentified digits. Accordingly, these trials were treated as if the participant had successfully detected all targets (i.e. before the over-reporting).

For explanation of choice of visual scoring method, see Appendix 3.1.

Test conditions (and trial structure): In total, there were 19 possible test conditions for stimulus presentation in the Digit Stream Dual-Task for the NH controls (Figure 3.1). These involved one condition for speech type (i.e. normal), three conditions for task type (i.e. visual task only, auditory task only, or dual-task), four conditions for noise type (i.e. quiet, or speech-shaped noise at 0dB, -3dB or -6dB SNR), and also four conditions for total number of digits presented within the target stream (i.e. 10, 20, 30 or 40 digits, 10 being the slowest). There were 20 trials for each test condition.

3.4. Experiment 1

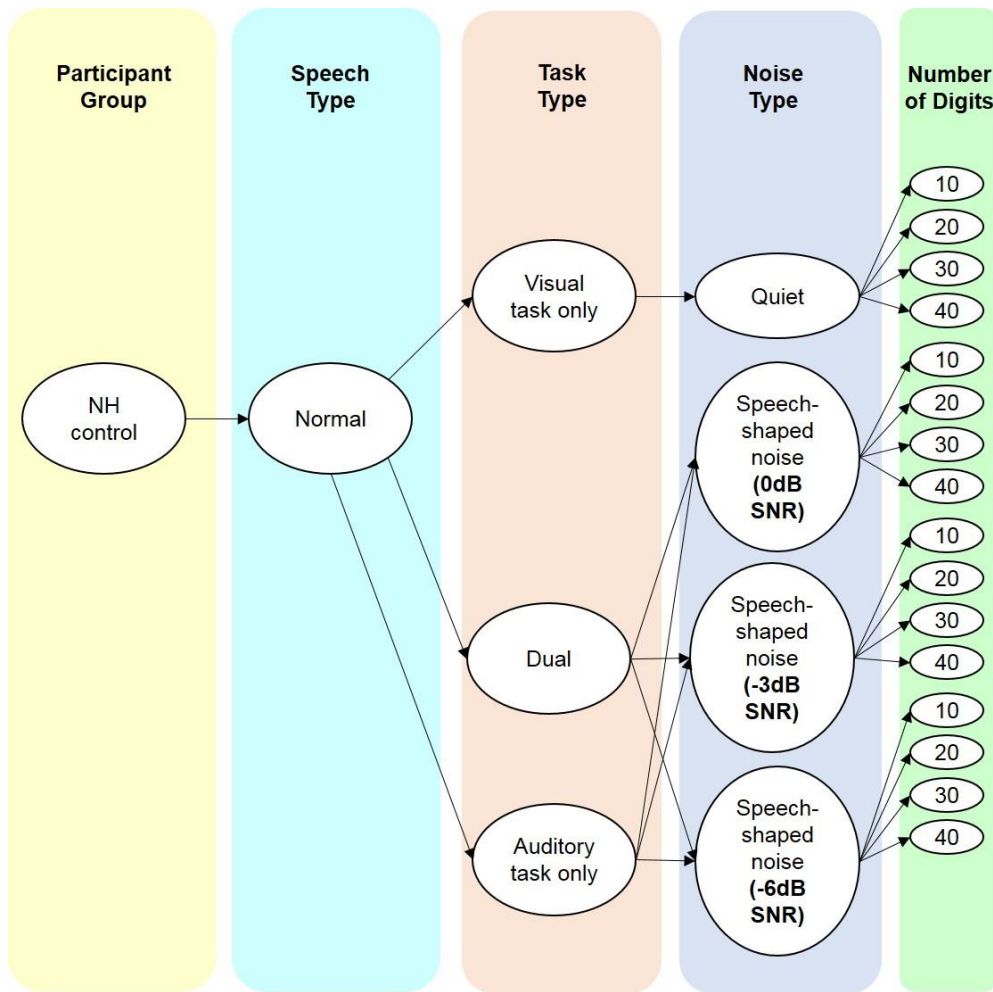


Figure 3.1: A schematic showing the breakdown of all test conditions within the Digit Stream Dual-Task for the NH controls (same structure as for Figure 2.2, but with the addition of number of digits being presented in the digit stream).

The duration of each trial was 3.2 seconds. Within the visual-only condition, only the digit stream was presented (Figure 3.2 a). Within the auditory-only condition, the GUI was blank, with the sentence utterance delivered after a 600 millisecond interval (Figure 3.2b). If speech-shaped noise was applicable to the test condition, the noise was present within this 600-millisecond interval before sentence onset, and also continued after sentence offset for the remainder of the trial before the response prompt.

For a dual-task condition, the digit stream began immediately in the GUI and was present within the 600-millisecond interval before auditory sentence (Figure 3.2 c). After sentence offset, the digit stream continued for the remainder of the trial before the response prompt. Two response boxes then appeared, first asking for the visual response and then the auditory response. When noise was applicable to the dual-task condition, the onset and offset of the noise were simultaneous with the onset and offset of digit stream.

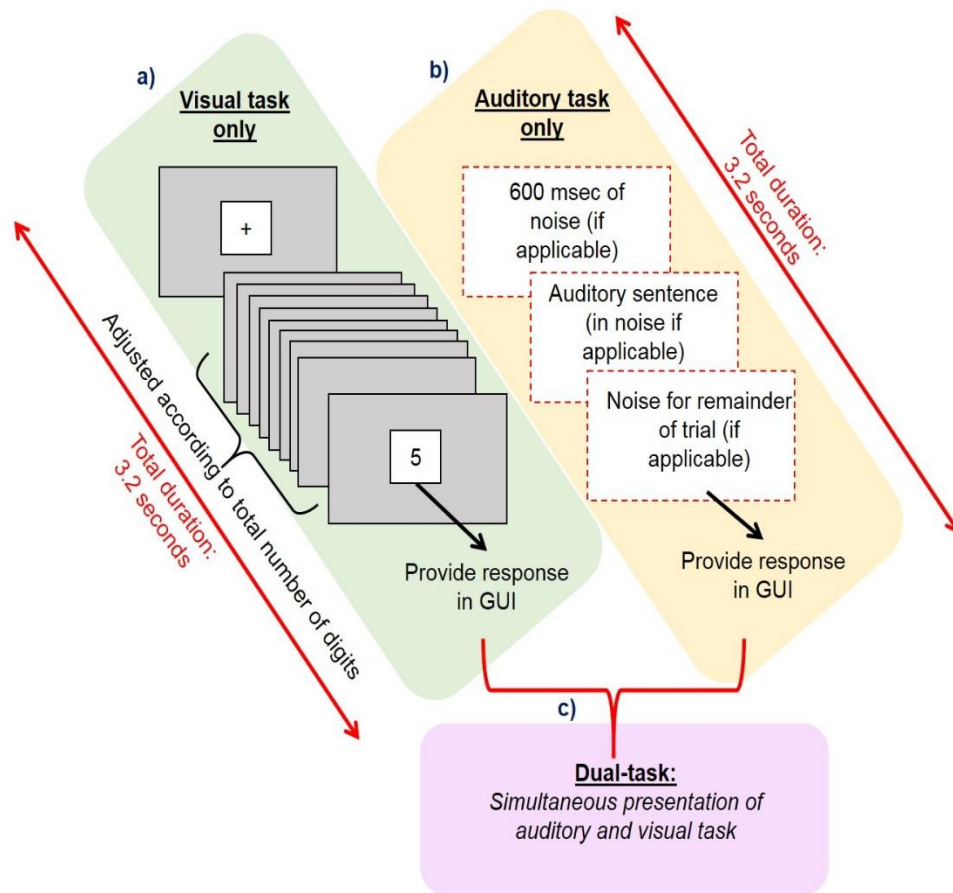


Figure 3.2: A schematic showing the structure of each trial for each task type in the test conditions. **(a)** Visual task only. **(b)** Auditory task only. **(c)** Dual-tasking.

It needs to be noted that, the faster the speed, the greater the number of targets presented. Accordingly, for the condition of 10 digits, there was a maximum of 2 targets; in the condition of 20 digits, the maximum target number was 5; for 30 digits, the maximum target number was 8; and for 40 digits, the maximum was 10 targets. There was always a target present in the stream (although the participant did not know this), and the target incidence frequency within a test condition expressed the full range of possibilities (i.e. from 1 target all the way to the maximum number of targets for that speed). Not only this, each possible target incidence frequency was also evenly distributed (as far as was possible) across all 20 trials.

Acoustic challenge: The same speech-shaped noise as used in the Pilot Study was applied, including a 100-millisecond raised cosine-smoothed onset.

Protocol: The testing session was divided into three experimental blocks: the “**Pre-Digit**” **Block**; then the “**Digit Stream Dual-Task**” **Block**; and finally the “**Post-Digit**” **Block**. The “Pre” and “Post” prefixes relate to the fact that the same test condition (i.e. the visual-only task) was run twice: once immediately before the participant was required to dual-task, and once again immediately after the participant completed dual-tasking. For the overall protocol of the testing session, see Table 3.1. Before the test conditions were implemented within the Digit Stream Dual-Task block, all participants underwent a practice run containing 5 trials each for the auditory task on its own, the visual task on its own (with 20 digits presented in total within the digit stream) and also the dual-task condition (again with 20 digits presented in total within the digit stream). The practice of the auditory task and the dual-task involved speech-shaped noise being presented at the easier SNR (i.e. 0dB SNR).

To control for fatigue and/or training effects, the order of the test conditions (i.e. steps 4-7 in the Pre-Digit Block, 1-15 in the Digit Stream Dual-Task Block, and steps 1-4 in the Post-Digit Block in Table 3.1) was randomised and counter-balanced. Furthermore, the lists of sentence stimuli were randomised and counterbalanced such that a participant never encountered the same sentence stimulus list twice within the entire testing session.

Statistical analysis: Due to the small sample size, only descriptive statistics were pursued.

3.4. Experiment 1

Table 3.1: Protocol for Experiment 1. Same detail provided as for Table 2.3, with the addition of block type (i.e. Pre-Digit, Digit Stream Dual-Task, or Post-Digit).

Block	Step	Training, Practice or Test?	Task Type or Test Condition	Speech Type	Noise Type	Total number of trials/ Total duration
Pre-Digit Block	1	Practice	Auditory task only	Normal	Speech-shaped noise at 0dB SNR	5 trials
	2	Practice	Visual task only at 20 digits	N/A	Quiet	5 trials
	3	Practice	Dual at 20 digits	Normal	Speech-shaped noise at 0dB SNR	5 trials
	4 *	Test	Visual task only at 10 digits	N/A	Quiet	20 trials
	5 *	Test	Visual task only at 20 digits	N/A	Quiet	20 trials
	6 *	Test	Visual task only at 30 digits	N/A	Quiet	20 trials
	7 *	Test	Visual task only at 40 digits	N/A	Quiet	20 trials
Digit Stream Dual-Task Block	1 *	Test	Dual at 10 digits	Normal	Speech-shaped noise at 0dB SNR	20 trials
	2 *	Test	Dual at 20 digits	Normal	Speech-shaped noise at 0dB SNR	20 trials
	3 *	Test	Dual at 30 digits	Normal	Speech-shaped noise at 0dB SNR	20 trials
	4 *	Test	Dual at 40 digits	Normal	Speech-shaped noise at 0dB SNR	20 trials
	5 *	Test	Dual at 10 digits	Normal	Speech-shaped noise at -3dB SNR	20 trials
	6 *	Test	Dual at 20 digits	Normal	Speech-shaped noise at -3dB SNR	20 trials
	7 *	Test	Dual at 30 digits	Normal	Speech-shaped noise at -3dB SNR	20 trials
	8 *	Test	Dual at 40 digits	Normal	Speech-shaped noise at -3dB SNR	20 trials
	9 *	Test	Dual at 10 digits	Normal	Speech-shaped noise at -6dB SNR	20 trials
	10 *	Test	Dual at 20 digits	Normal	Speech-shaped noise at -6dB SNR	20 trials
	11 *	Test	Dual at 30 digits	Normal	Speech-shaped noise at -6dB SNR	20 trials
	12 *	Test	Dual at 40 digits	Normal	Speech-shaped noise at -6dB SNR	20 trials
	13 *	Test	Auditory task only	Normal	Speech-shaped noise at 0dB SNR	20 trials
	14 *	Test	Auditory task only	Normal	Speech-shaped noise at -3dB SNR	20 trials
	15 *	Test	Auditory task only	Normal	Speech-shaped noise at -6dB SNR	20 trials
Post-Digit Block	1 *	Test	Visual task only at 10 digits	N/A	Quiet	20 trials
	2 *	Test	Visual task only at 20 digits	N/A	Quiet	20 trials
	3 *	Test	Visual task only at 30 digits	N/A	Quiet	20 trials
	4 *	Test	Visual task only at 40 digits	N/A	Quiet	20 trials

3.4.3. Results and Discussion: Descriptive statistics

3.4.3.1. Secondary task performance: Visual accuracy

At the group level, visual accuracy appeared to deteriorate in each of the digit stream speeds when the participant was required to dual-task (compared to performing just the visual task), with further deterioration when SNR worsened (Figure 3.3 a, b, c and d). This was as originally hypothesised. However, this trend also needed to apply to individual participant performances. In order to ascertain this, an average was taken of the two visual-only conditions for each condition of digit stream speed (i.e. the visual-only condition in the Pre-Digit block and the visual-only condition in the Post-Digit block). The target detection percentage in each of the dual-tasking conditions (that were executed during the Digit Stream Dual-Task block) were then subtracted from this average (within each digit stream speed). The subtraction was such that, if a positive value was elicited, this indicated deterioration in visual accuracy when the participant was dual-tasking relative to executing only the visual task (it needs to be noted that this subtraction does not amount to a calculation of LE, because no comparison against dual-tasking in quiet was possible).

It was revealed that, as the digit stream speed increased, the NH controls tended to elicit a negative difference in visual accuracy between single-task and dual-task conditions (Figure 3.3 e, f, g and h), i.e. they were more accurate when dual-tasking than in the visual-only condition. Furthermore, there was relatively little change in visual performance between the different SNRs at a given digit stream speed. Therefore, individual participant performances were not in keeping with the hypotheses. This gives rise to the possibility that inappropriate resource allocation could have occurred during the dual-tasking, i.e. more cognitive resources were assigned to the visual task, instead of the primary task. If the secondary task has indeed taken precedence over the primary task, then this would compromise any future LE measurement conducted with this type of digit stream secondary task.

3.4. Experiment 1

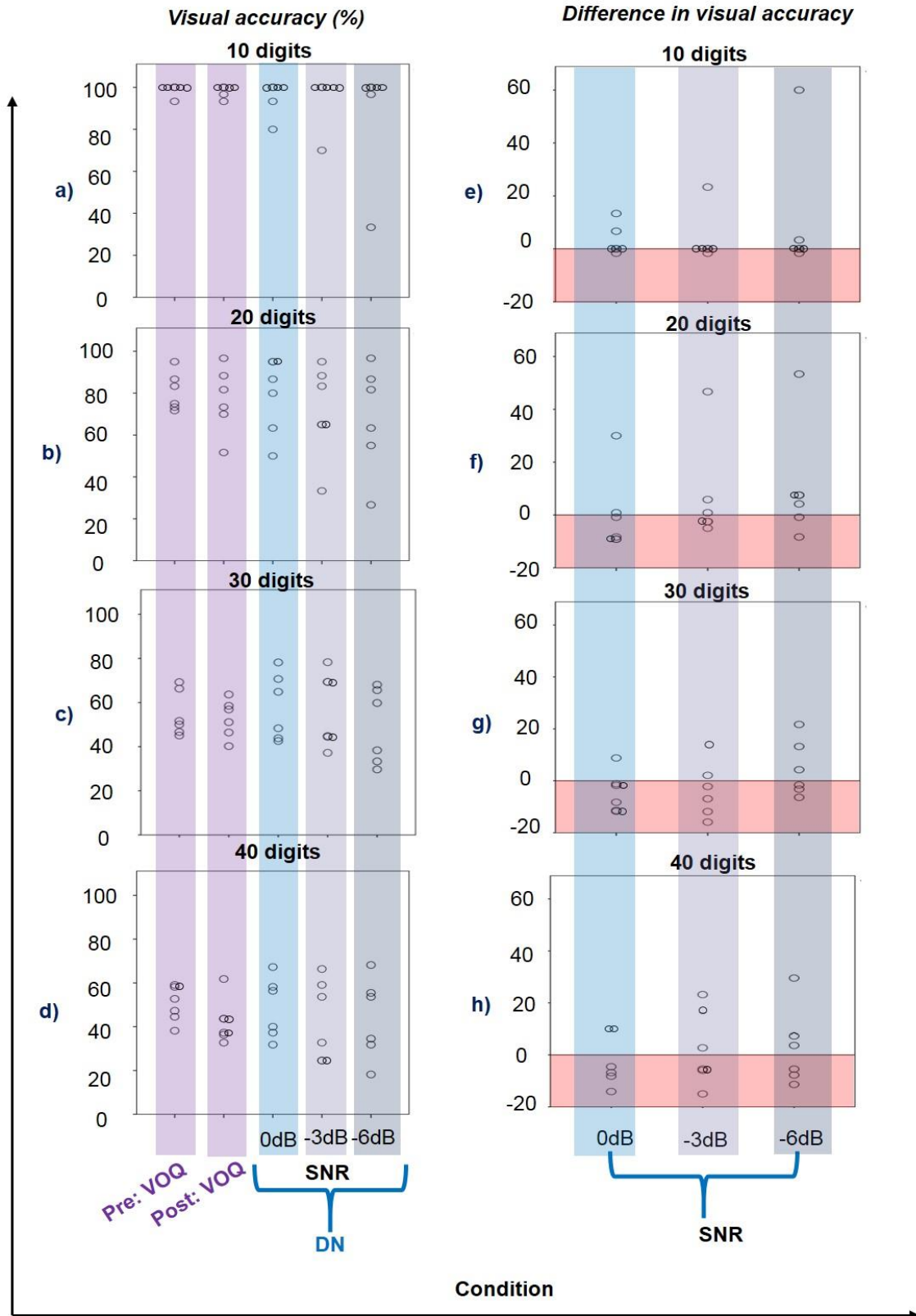


Figure 3.3: Dot plots of visual accuracy (a, b, c, d) and difference of visual performance between dual-task and single-task conditions (e, f, g, h) across the different stream speeds (unit = difference in percentage).

Pre: VOQ = Visual task only (in quiet) in the Pre-Digit Block, Post: VOQ = Visual task only (in quiet) in the Post-Digit Block, DN = Dual-tasking in noise. Red shading indicates when the participants' performance improved within the dual-task condition compared to performing just the visual task on its own.

3.4. Experiment 1

3.4.3.2. Stability of primary task performance: Auditory accuracy

To determine the likelihood that inappropriate resource allocation had occurred, primary task stability needs to be determined. It emerged that, as SNR worsened, auditory accuracy decreased within both the single-task and dual-task conditions (Figure 3.4 a). However, it was also revealed that across the test conditions of digit stream speed and SNR, half of the NH controls exhibited deterioration in auditory performance when dual-tasking (Figure 3.4 b).

It does need to be noted that, in general, the NH controls exhibited variability in auditory accuracy. This could arise from measurement variability. Nevertheless, this reflects the same issue previously reported in the Pilot Study, i.e. the concern that the occurrence of primary task instability cannot be discounted and, therefore, the participants may have executed inappropriate resource allocation.

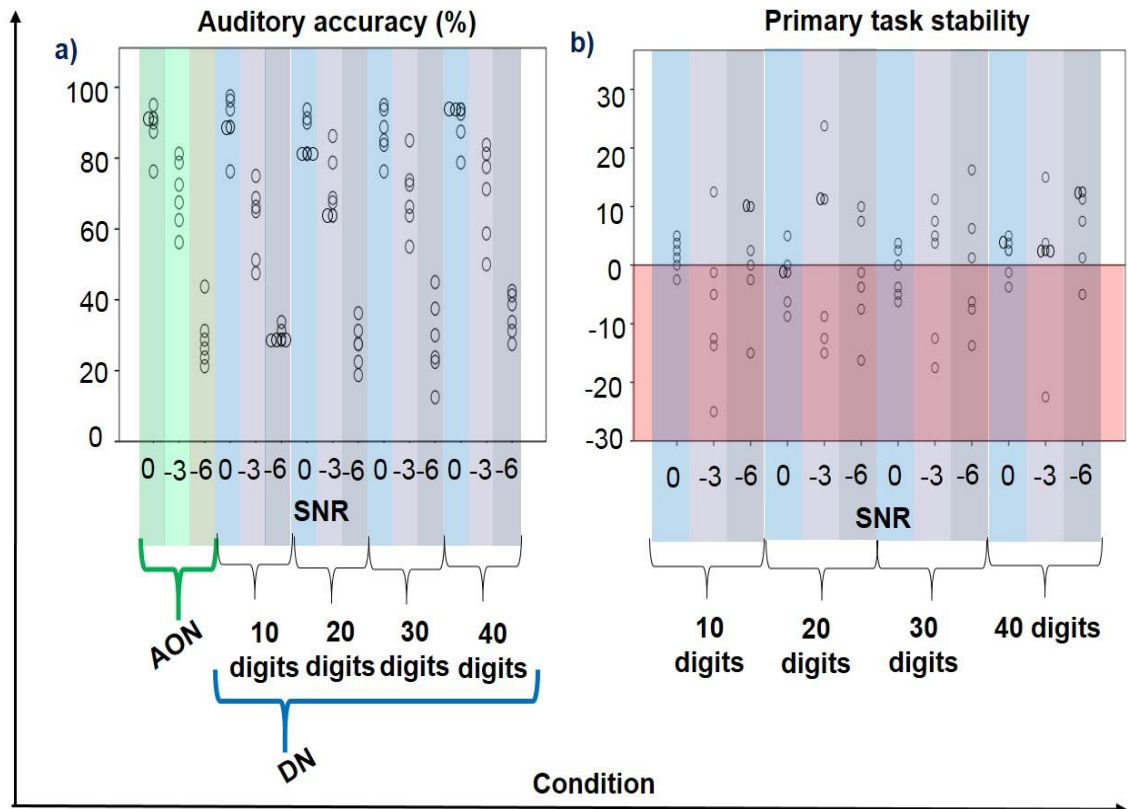


Figure 3.4: Dot plots of auditory accuracy **(a)** and primary task stability **(b)** across different conditions of SNR and digit stream speed.

AON = Auditory task only in noise, DN = Dual-tasking in noise. Red shading indicates deterioration in primary task performance when dual-tasking.

Unit of primary task stability = Difference in percentage.

3.4.3.3. *Evaluation of potential viability of new LE measurement*

To investigate whether there was any aspect of the current digit stream that was possibly causing this incorrect prioritisation in cognitive resources, the visual accuracy data was revisited. Specifically, the trials where over-reporting had occurred were tallied up in all test conditions for each target presentation frequency (i.e. 1 target all the way to 10 targets: Figure 3.5).

Over-reporting occurred when there were 1-4 targets presented within the digit stream, with a maximum of 6 trials identified for a participant within a given test condition. The occurrence of over-reporting decreased as the number of targets increased. This reduction reached the point that, when 4 targets were presented, only one participant was over-reporting within a single trial during the Pre-Digit block. However, everyone was culpable of over-reporting, and none of the test conditions avoided this phenomenon. This consistency in occurrence gives rise to the possibility that there is some kind of process (physiological or psychological) that might be interfering with the perception and, therefore, performance of the secondary task.

Therefore, it seems that the reliability of secondary task performance has become compromised, thus bringing the viability of this digit stream dual-task as a form of LE measurement into question.

3.4. Experiment 1

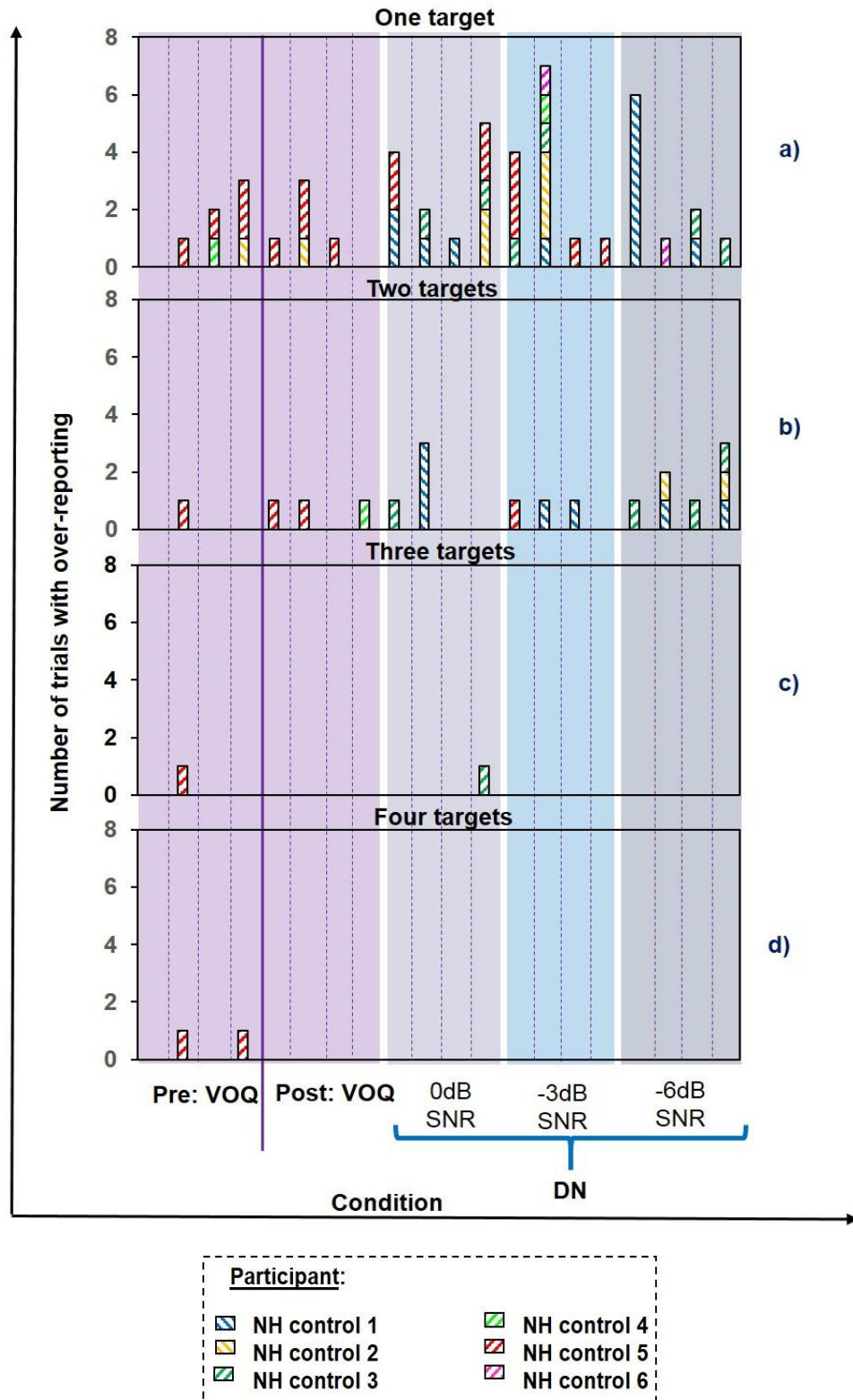


Figure 3.5: Stacked histograms showing the incidence of over-reporting for each participant for target presentation frequencies of 1-4.

Each coloured column is subdivided into 4 (by vertical dashed lines) to represent different digit stream speeds (from left to right: speed produced by having a total of 10, 20, 30 and 40 digits within stream).

Pre: VOQ = Visual task only in the Pre-Digit Block, Post: VOQ = Visual task only in the Post-Digit Block, DN = Dual-tasking in noise.

3.5. Conclusions of Experiment 1

It is not possible to ignore the notion that the NH controls may have used inappropriate resource allocation when dual-tasking, especially in light of the potential primary task instability. Furthermore, the participants demonstrated a concerning tendency to report more targets than actually were presented. This over-reporting could be potentially symptomatic of an underlying physiological, or psychological, process. Therefore, unless this phenomenon of over-reporting can be ameliorated, the capacity of the digit stream dual-task for future LE measurement is compromised.

Therefore, within Experiment 2, there will be an attempt to remove this over-reporting phenomenon.

3.6. Proposed solution

A possible cause of the over-reporting phenomenon is retinal image persistence. This is where the physical stimulus presentation has become quicker than the physiological process of transduction at the retina, meaning that the perception of a specific stimulus actually lingers after it had been removed from view (i.e. a positive after-image), which would then skew perception (Coltheart, 1980; Shimojo et al., 2001). The exact cause of this lingering after-image is not well established, but one postulated mechanism is that there is a persistence in neuronal impulses being sent to the occipital cortex from the retina, potentially due to the bleaching of photochemical pigments and/or adaptation within the neuronal networks at the retina (Coltheart, 1980; Shimojo et al., 2001). This after-image is typically brief in duration (less than 500 milliseconds: Coltheart, 1980; Shimojo et al., 2001). However, due to the rapid nature of visual presentation within the digit stream, this brevity is unlikely to minimise its impact on the participant's response if retinal image persistence is indeed occurring.

A potential approach to prevent (or at least minimise) retinal image persistence is the application of backward masking (Breitmeyer, 2007). Visual backward masking involves the flashing of a very brief (≤ 50 milliseconds) visual stimulus (i.e. the mask) immediately after the visual item, thus interrupting the processing of the just presented visual item (Breitmeyer, 2007; Kinsbourne & Warrington, 1962). This technique has long been applied within visual perceptual research (being first attempted in the late 19th century) as well as within the investigation of consciousness (Breitmeyer & Ögmen, 2006; Sherrington, 1897; Vorberg et al., 2003; Breitmeyer & Ögmen, 2004).

Therefore, Experiment 2 will incorporate this masking technique. However, the integration of masks within the digit stream will inevitably change the testing parameters of the secondary task. Therefore, another small-scale study (entailing performance of only the secondary task) is required, in order to choose a digit stream speed, before this modified digit stream dual-task can be trialled as a form of LE measurement.

3.7. Experiment 2

3.7.1. Hypotheses

- As digit stream speed increases, visual accuracy (i.e. target detection percentage) will decrease.
- The introduction of backward masking will reduce the occurrence of over-reporting.

3.7.2. Methods

3.7.2.1. Participants

Two of the original (NH) controls (1 female and 1 male) were tested again in this experiment, and two new NH controls were recruited (1 female and 1 male). All participants had normal, or corrected-to-normal vision, and were native British English speakers (mean age: 55.2 years, \pm 17.3 S.D).

3.7.2.2. Procedure

Location of testing: Testing sessions were conducted within the participants' homes.

Secondary task of the Digit Stream Dual-Task: The same as for Experiment 1, but with the following modification: the introduction of a random noise visual mask after each digit (for close-up, see Figure 3.7). This mask comprised of a collection of small squares that were randomly black or white and each lasted for 50 milliseconds (in accordance with the backward masking technique: Kinsbourne & Warrington, 1962; Breitmeyer, 2007).

Test conditions (and trial structure): Seven different digit stream speeds were tested, which were created by changing the total number of digits presented in the stream (i.e. 10, 15, 20, 25, 30, 35 and 40 digits – see Figure 3.6). There were 20 trials for each test condition. Before the test conditions were implemented, all participants underwent a practice run where they were able to execute 5 trials with the total number of digits presented in the stream fixed at 20 digits.

3.7. Experiment 2

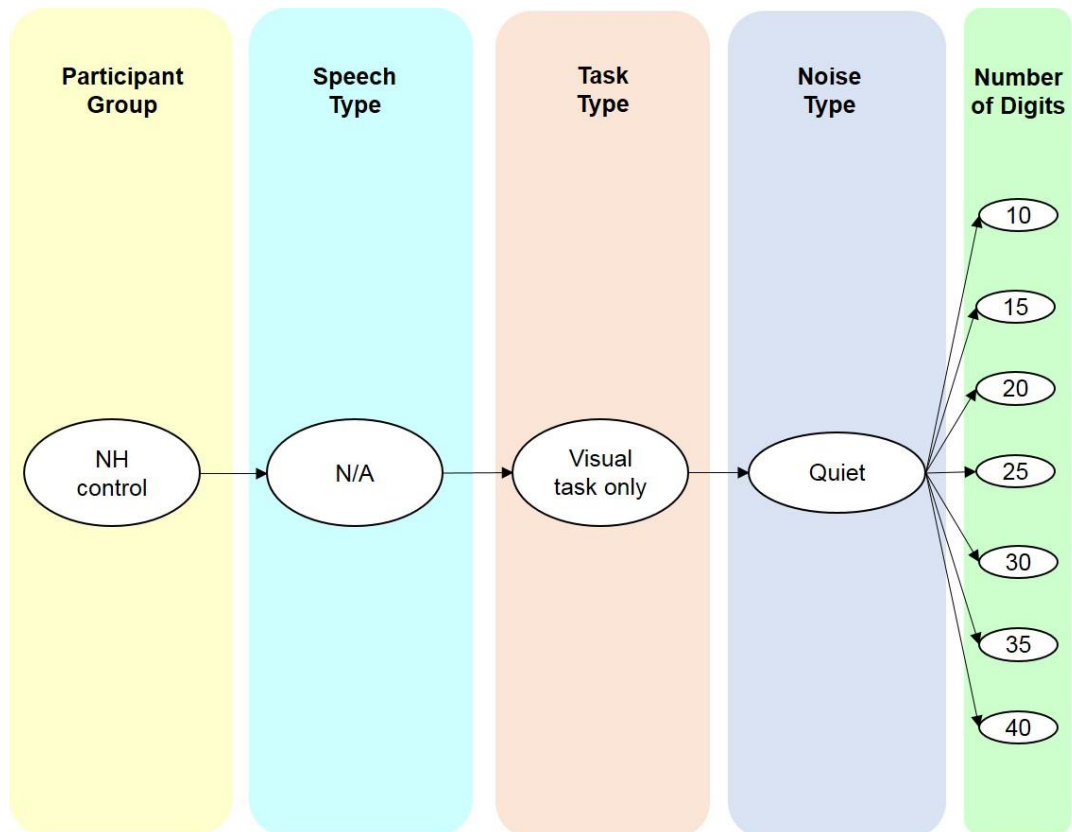


Figure 3.6: A schematic of test conditions (same structure as for Figure 3.1, but with the addition of number of digits being presented in the digit stream).

The trial structure was the same as for the visual-only condition presented within Experiment 1 (Figure 3.7).

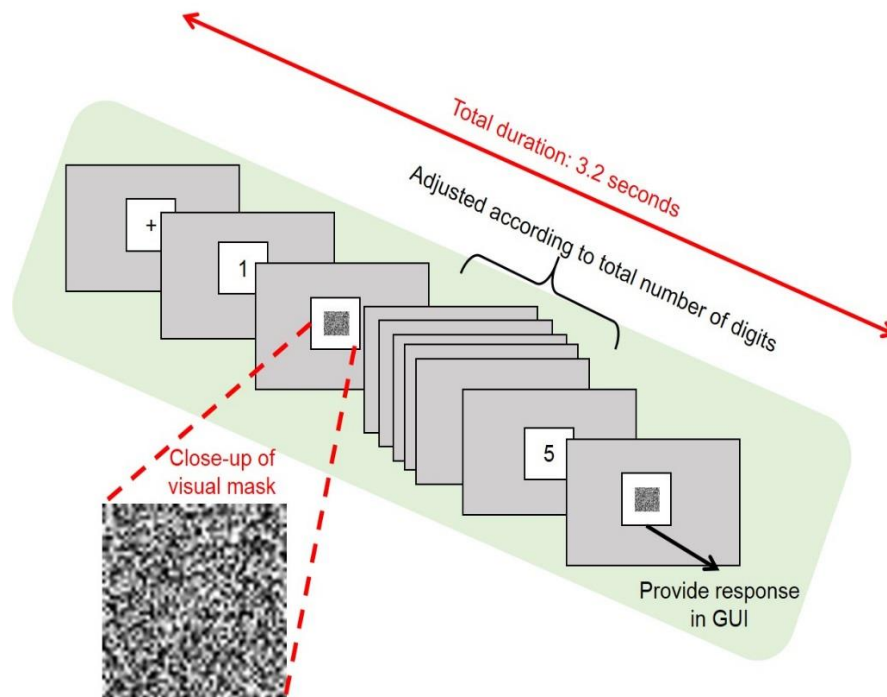


Figure 3.7: A schematic showing trial structure, with a close-up of the random noise visual mask also provided.

3.7. Experiment 2

When the total of digits in a stream was 10 and 15, the maximum number of targets that could be presented was 2. This maximum changed to 3 and 4 when the total digits presented became 20 and 25 digits respectively. When there was 30, 35 or 40 digits in total within the stream, the maximum number of targets increased to 5. In each test condition, the target incidence frequency expressed the full range of possibilities (i.e. from 1 target all the way to the maximum number of targets), with each possible incidence frequency being evenly distributed (as far as was possible) across all 20 trials.

Protocol: For the overall protocol of the testing session, see Table 3.2. To control for fatigue and/or training effects, the order of the test conditions (i.e. steps 2-8 in Table 3.2) were randomised and counter-balanced.

Table 3.2: Protocol for Experiment 2. Same detail provided as for Table 2.3, including red numbers (with asterisks) indicating all the steps in the protocol where the order of the test conditions were randomised and counter-balanced together.

Step	Training, Practice or Test?	Task Type or Test Condition	Speech Type	Noise Type	Total number of trials/ Total duration
1	<i>Practice</i>	Visual task only at 20 digits	N/A	Quiet	5 trials
2 *	Test	Visual task only at 10 digits	N/A	Quiet	20 trials
3 *	Test	Visual task only at 15 digits	N/A	Quiet	20 trials
4 *	Test	Visual task only at 20 digits	N/A	Quiet	20 trials
5 *	Test	Visual task only at 25 digits	N/A	Quiet	20 trials
6 *	Test	Visual task only at 30 digits	N/A	Quiet	20 trials
7 *	Test	Visual task only at 35 digits	N/A	Quiet	20 trials
8 *	Test	Visual task only at 40 digits	N/A	Quiet	20 trials

Statistical analysis: Due to the small sample size, only descriptive statistics were pursued.

3.7.3. Results and Discussion: Descriptive statistics

3.7.3.1. Secondary task performance: Visual accuracy

Overall, when the incidence of over-reporting was assessed in Experiment 2, a maximum of only 2 trials was contaminated by over-reporting for a participant within a given test condition (Figure 3.8 a, b, c, and d). Over-reporting was non-existent when there were 5 targets presented in the digit stream (applicable to digit stream speeds generated by having 30, 35 and 40 digits within the stream).

Furthermore, the scarcity in over-reporting applied to the naïve participants (NH controls 3 and 4) as well as the original two participants from Experiment 1 (NH controls 1 and 2).

The visual accuracy scores were next examined to ascertain whether there was an ideal digit stream speed for use within the digit stream dual-task (i.e. to optimise sensitivity in LE measurement – see Figure 3.8 e). The speed produced by having a total of 10 digits in the stream led to ceiling performance of 100% detection in all participants. On the other hand, when the total number of digits presented increased to 30, 35 and 40, detection dropped beneath 60% for nearly all NH controls.

The speed generated by the total of 25 digits engendered a fairly evenly dispersed performance across participants, with 100% detection being possible but not achieved by everyone. Therefore, this speed seems to be a prime candidate to use when the digit stream is integrated within the dual-task framework for LE measurement. This is because this particular speed appears to offer enough manoeuvrability in the scores to be able to differentiate between high-level and relatively mediocre performers. Furthermore, performance never deteriorated beyond 70%, meaning that there is also margin for the detection of particularly poor performers in future testing (because average performance is not already at floor level).

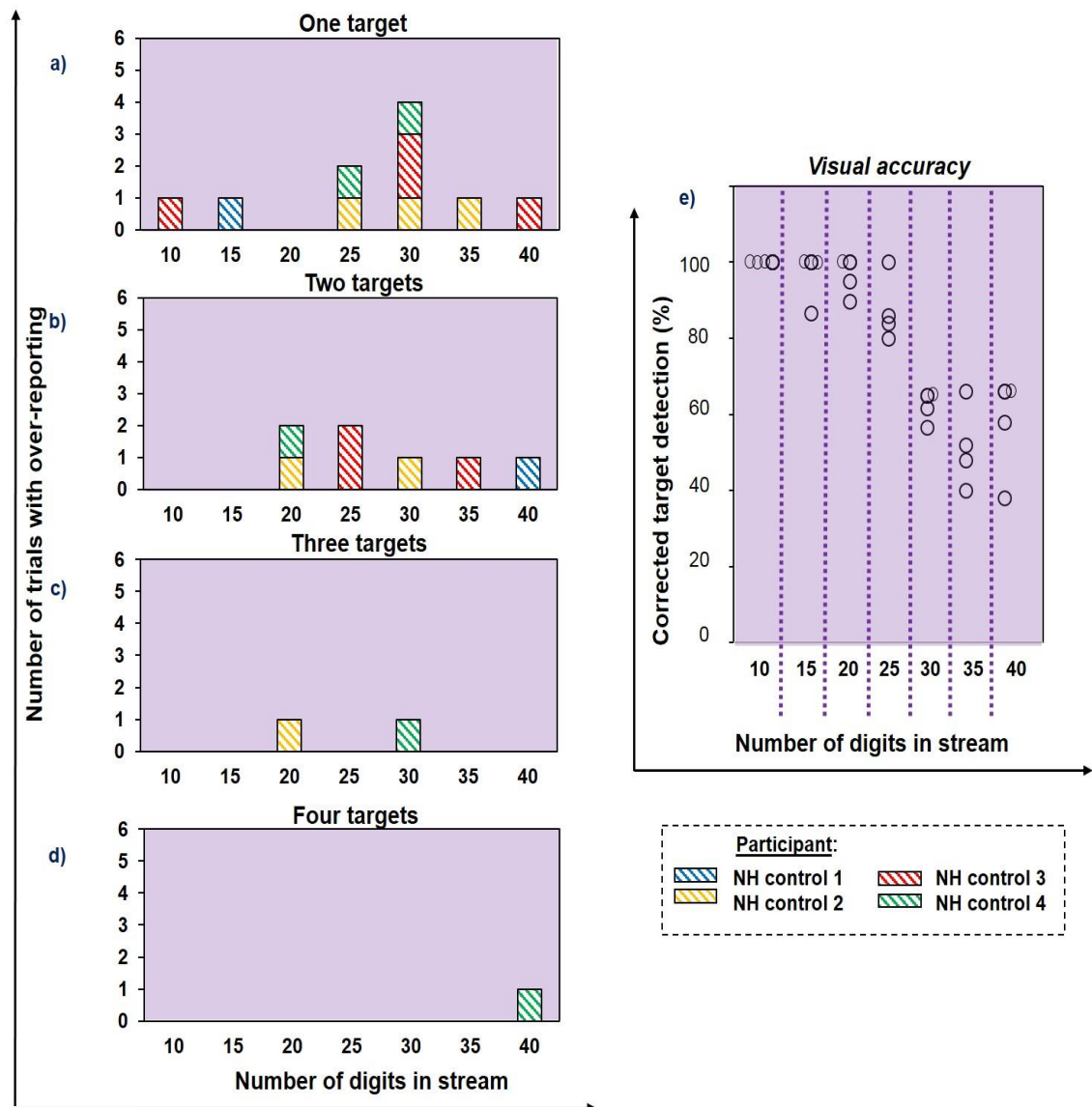


Figure 3.8: Stacked histograms showing the incidence of over-reporting as a function of digit stream speed at different target presentation frequencies (**a, b, c, d**) and dot plot of visual accuracy for all NH controls, with target detection percentage being corrected for all instances of over-reporting (**e**).

3.8. Conclusions of Experiment 2

The speed produced by having a total of 25 digits presented in the stream (with masks interspersed) appears to possess high granularity for scoring visual accuracy. Also, the introduction of the mask provides more reassurance that the phenomenon of retinal image persistence has been addressed, and should no longer be a primary driver of any over-reporting that occurs. However, it is noticeable that over-reporting did still persist, although the incidence of this appears to be relatively low.

3.9. Experiment 3

3.9.1. Introduction

It is now necessary to test this refined digit stream dual-task properly with a larger cohort of participants to enable inferential statistical analysis and the determination of the potential viability of the digit stream dual-task as a LE test. Experiment 3 will, therefore, be the first larger-scale study of this thesis. Subjective ratings will also be concurrently collected.

3.9.2. Hypotheses

When there is acoustic challenge, the digit stream dual-task will show the following:

- Deterioration in secondary task performance (i.e. positive LE scores will be attained).

This increase in LE will not be at the cost of primary task stability (i.e. resource allocation is as intended):

- Auditory accuracy will remain constant when the participant is dual-tasking (relative to that achieved when performing only the auditory task).

In addition, there will be subjective corroboration of the behavioural data:

- Perceived effort will increase as LE scores increase.

3.9.3. Methods

3.9.3.1. Participants

Twenty eight adult normal hearing (NH) controls (12 females and 16 males, mean age: 45.8 years, \pm 12.5 S.D) were recruited. All participants had normal, or corrected-to-normal vision, and were native British English speakers.

3.9.3.2. Procedure

Location of testing: Testing sessions were conducted in the participants' office environment (within a quiet conference room).

Primary task of the Digit Stream Dual-Task: The same as for Experiment 1.

Secondary task of the Digit Stream Dual-Task: The same as for Experiment 2, except that the digit stream speed was now fixed, with a total of 25 digits presented in each stream (with masks interspersed). The maximum number of targets that could be presented within a given stream was 6.

Acoustic challenge: The same as for Experiment 1.

Other LE indices: Subjective ratings were conducted using the same 5-point Likert scale as implemented in the Pilot Study, with the following modification: for test conditions that involved dual-tasking, participants were required to rate the perceived effort for the overall dual-task (instead of providing individual ratings for the primary and secondary task components as in the Pilot Study).

Test conditions (and trial structure): In total, there were six possible test conditions (Figure 3.9), of which there was one condition for speech type (i.e. normal), as well as three conditions for task type (i.e. visual task only, auditory task only, or dual-task) and three conditions for noise type (i.e. quiet, speech-shaped noise at 0dB SNR, or speech-shaped noise at -6dB SNR). There were 18 trials for each test condition. The trial structure was consistent with that utilised in Experiment 1.

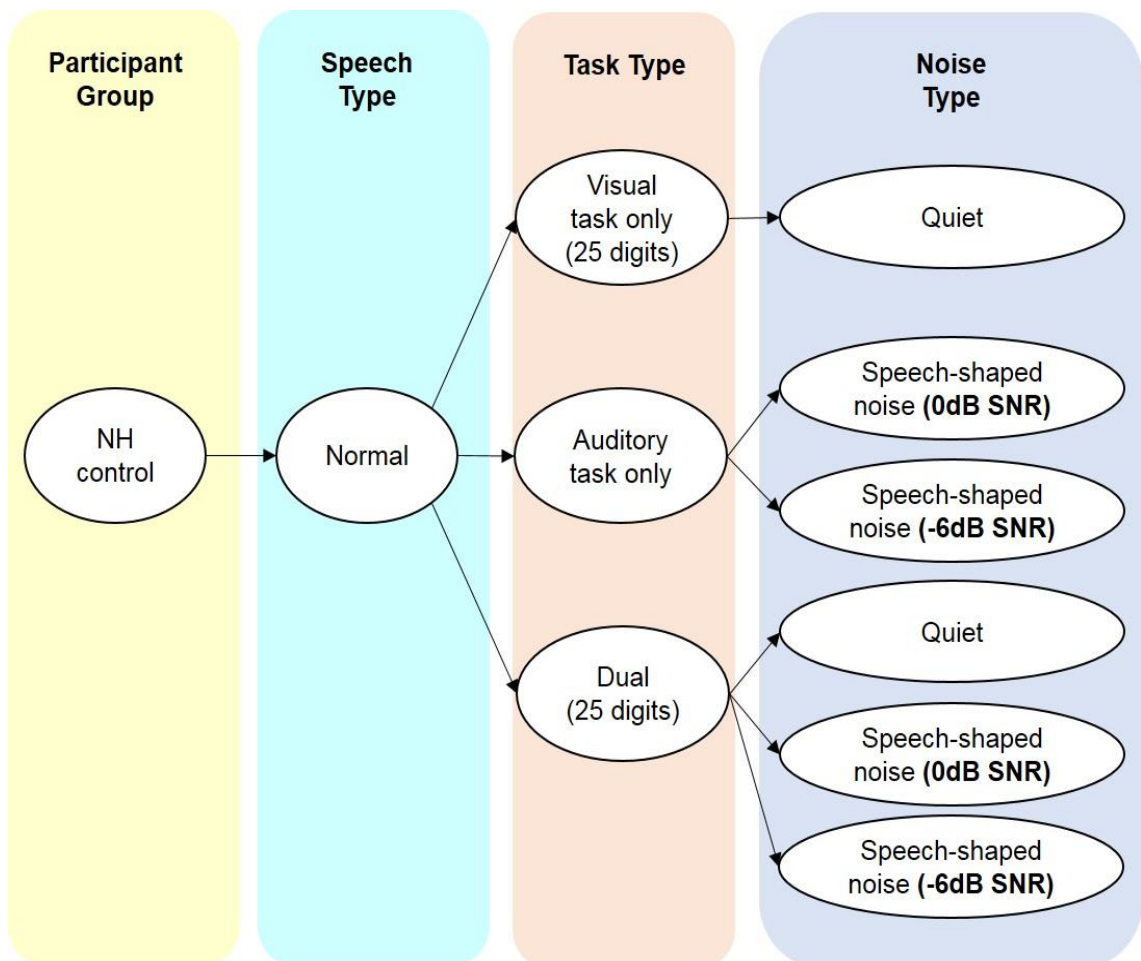


Figure 3.9: A schematic of test conditions (same structure as for Figure 3.1).

Protocol: For the overall protocol, see Table 3.3. Before the test conditions were executed, all participants underwent a practice run where they were able to execute 5 trials for the auditory task on its own (conducted in quiet) and also for the dual-task condition (again conducted in quiet). To control for fatigue and/or training effects, the order of the test conditions (i.e. steps 3-8 in Table 3.3) were randomised and counter-balanced.

Table 3.3: Protocol for Experiment 3. Same detail provided as for Table 2.3, including red numbers (with asterisks) indicating all the steps in the protocol where the order of the test conditions were randomised and counter-balanced together.

Step	Training, Practice or Test?	Task Type	Speech Type	Noise Type	Total number of trials/ Total duration
1	<i>Practice</i>	Auditory task only	Normal	Quiet	5 trials
2	<i>Practice</i>	Dual at 25 digits	Normal	Quiet	5 trials
3 *	Test	Visual task only at 25 digits	N/A	Quiet	18 trials
4 *	Test	Auditory task only	Normal	Speech-shaped noise at 0dB SNR	18 trials
5 *	Test	Auditory task only	Normal	Speech-shaped noise at -6dB SNR	18 trials
6 *	Test	Dual at 25 digits	Normal	Quiet	18 trials
7 *	Test	Dual at 25 digits	Normal	Speech-shaped noise at -6dB SNR	18 trials
8 *	Test	Dual at 25 digits	Normal	Speech-shaped noise at 0dB SNR	18 trials

Statistical analysis: SPSS was utilised to execute mixed effects logistic regression and a mixed effects linear model (SPSS Statistics 22, IBM).

3.9.4. Results and Discussion: Descriptive and inferential statistics

3.9.4.1. *LE measurement (secondary task performance): Visual accuracy*

Overall, there was relatively little change in target detection accuracy between performing just the visual task and dual-tasking in quiet (Figure 3.10 a). When noise was introduced, visual accuracy deteriorated. There was considerable variability in visual performance across all test conditions, even when it was just the visual task being performed (scores ranged from just above 60% all the way to above 90%). This variability appears to be intensified at the hardest noise level, with scores ranging from 40% to above 90%.

The majority of NH controls obtained positive LE scores (as calculated by difference in visual accuracy percentage) in both noise conditions (Figure 3.10 b), indicating that acoustic challenge did appear to increase the level of LE (as originally hypothesised). In the quiet listening conditions, there appeared to be little change in LE (which was also as expected). However, several NH controls obtained negative LE scores in all listening conditions, which was counter-intuitive.

To assess whether or not these trends in visual accuracy across conditions were of statistical significance, two models were generated. The first examined the degree of LE in quiet, by comparing visual accuracy when performing the visual task only, and when dual-tasking in quiet. This involved a mixed effects logistic regression model, with a single predictor of test condition. Participants were set as random effects, as part of the attempt to generalise the findings to the population. No statistically significant results were produced ($F(1,54)=2.5$, $p=.12$). This thus suggested that LE in quiet was not substantial.

The second model investigated LE exerted in noise, by comparing visual accuracy when dual-tasking in quiet and when dual-tasking in each noise level. Another mixed effects logistic regression model was applied, with a single predictor of test condition. Participants were again set as random effects. A statistically significant result ($F(2, 81)=15.6$, $p<.001$) suggested that visual accuracy differed across the three test conditions. Therefore, post-hoc pairwise contrasts (Bonferroni-corrected) were pursued on the mean visual accuracy in each condition (Table 3.4). These comparisons revealed that LE significantly increased at both SNRs (0dB: $p=.001$, -6dB: $p<.001$). However, the 6dB difference in SNR itself did not significantly augment the LE level ($p<.09$), i.e. the visual performance achieved when dual-tasking at the two noise levels did not significantly differ from each other.

3.9. Experiment 3

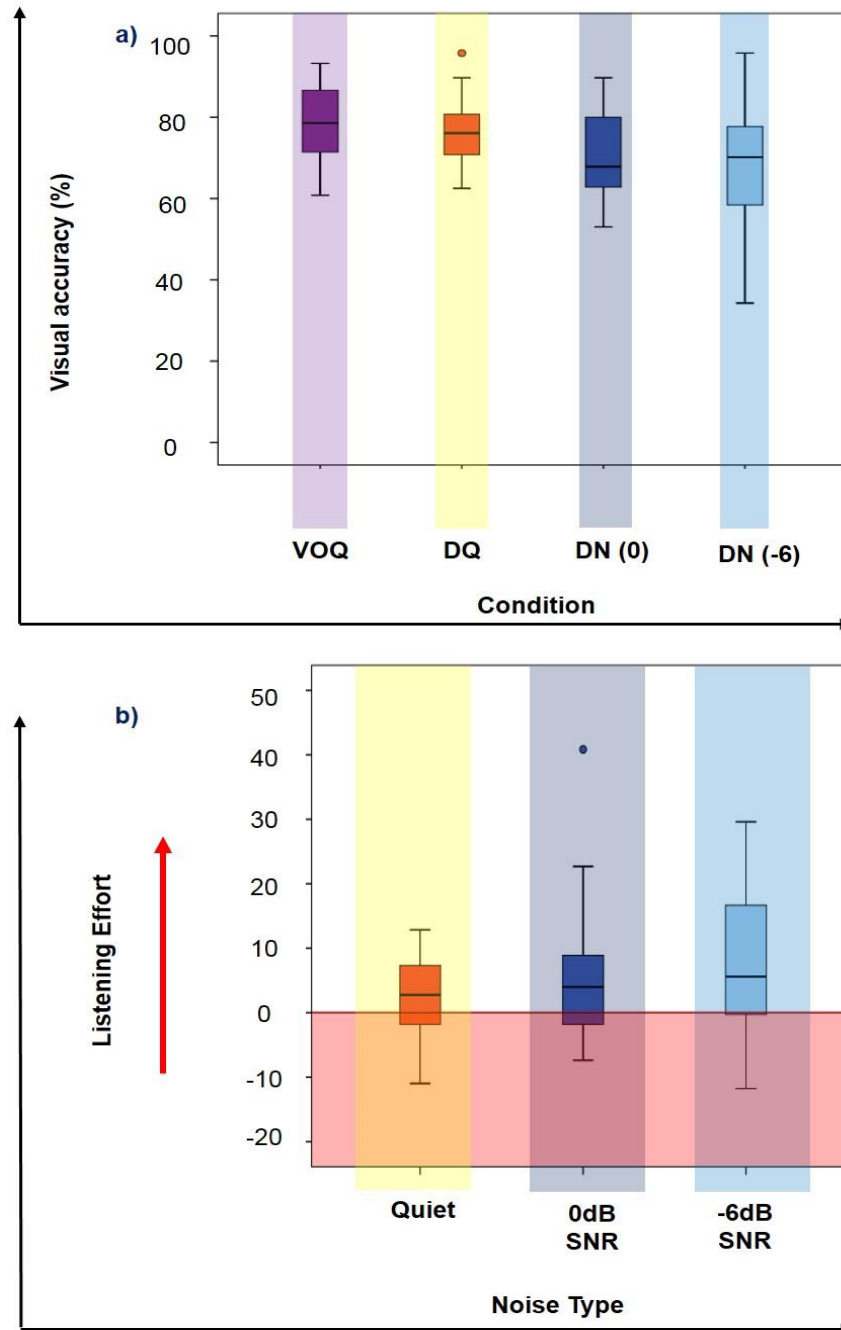


Figure 3.10: Boxplots of visual accuracy (a) and LE (b).

VOQ = Visual task only in quiet, DQ = Dual-tasking in quiet, DN (0) = Dual-tasking in noise of 0dB SNR, DN (-6) = Dual-tasking in noise of -6dB SNR. Red arrow indicates direction in which LE increases. Red shading indicates when LE has decreased in presence of noise (unit = difference in percentage).

Table 3.4: Mean and standard error of visual accuracy for each test condition

Test condition	Mean	Standard error
Visual only in quiet	78.5	1.6
Dual-tasking in quiet	76.4	1.5
Dual-tasking at 0dB SNR	71.2	2.1
Dual-tasking at -6dB SNR	68.7	2.7

3.9.4.2. Evaluation of validity of LE measurement: Subjective ratings

For the majority of NH controls, the perceived effort of performing just the visual task on its own was in the region of “medium hard work”. This also applied to dual-tasking in quiet. There was then a systematic shift in ratings from “medium hard work” towards that of “very hard work” when dual-tasking in noise at 0dB SNR, and “extremely hard work” when dual-tasking at -6dB SNR. However, similarly to that observed within the visual accuracy scores, there was considerable variability in the subjective ratings provided by the NH controls in all conditions (Figure 3.11). Therefore, the statistical significance of the observed trend was further assessed.

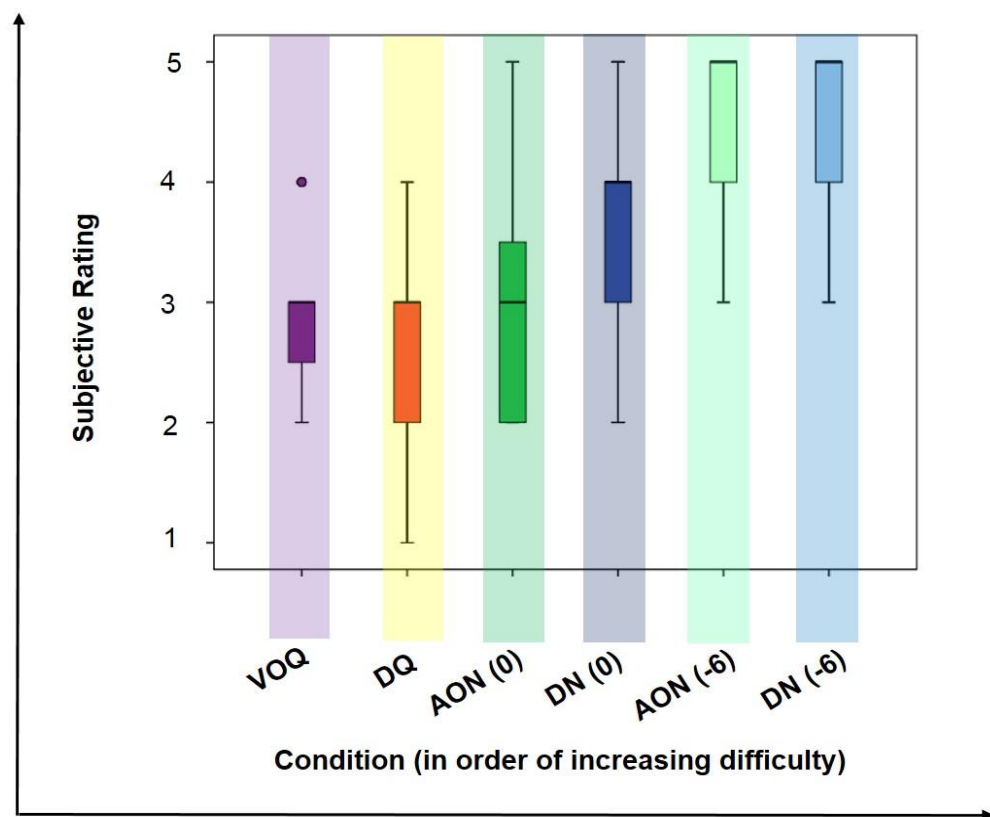


Figure 3.11: Boxplots of subjective ratings.

VOQ = Visual task only in quiet, DQ = Dual-tasking in quiet, AON (0) = Auditory task only in noise of 0dB SNR, DN (0) = Dual-tasking in noise of 0dB SNR, AON (-6) = Auditory task only in noise of -6dB SNR, DN (-6) = Dual-tasking in noise of -6dB SNR.

Subjective ratings: 1 = Not at all hard work, 2 = Quite hard work, 3 = Medium hard work, 4 = Very hard work, 5 = Extremely hard work.

Ideally, multinomial ordered logistic regression should be applied. However, this analytical technique is extremely complex. Therefore, for simplicity, the subjective ratings will be treated as a continuous measure. This will then enable the use of a mixed effects linear model.

3.9. Experiment 3

While acknowledging that this is far from an ideal modelling approach, the linear model does preserve the ordered nature of the subjective measure (i.e. the ratings lie on a scale of 1-5). This is important because, whilst these ratings comprise a categorical variable, the meaning lies in their order (a lower value indicates low perceived effort whilst a higher value indicates greater perceived effort).

This mixed effects linear model had a single predictor of test condition (with six levels). Participants were set as random effects. A statistically significant result ($F(5, 135)=60.5$, $p<.001$) suggests that subjective ratings differed across test condition. Therefore, post-hoc pairwise comparisons (Bonferroni-corrected) were pursued on the mean subjective rating in each condition (Table 3.5).

Table 3.5: Mean and standard error of subjective ratings for each test condition

Test condition	Mean	Standard error
Visual only	3.0	0.1
Dual-tasking in quiet	2.5	0.2
Auditory only at 0dB SNR	3.0	0.2
Dual-tasking at 0dB SNR	3.7	0.1
Auditory only at -6dB SNR	4.6	0.1
Dual-tasking at -6dB SNR	4.6	0.1

Only three comparisons were found to be insignificant:

- Auditory only at 0dB SNR versus visual only ($p=1$)
- Visual only versus dual-tasking in quiet ($p=.07$)
- Dual-tasking at -6dB SNR versus auditory only at -6dB SNR ($p=1$)

Otherwise, all comparisons were of significance ($p<.001$, except for the comparison between auditory only at 0dB SNR and dual-tasking in quiet, which yielded $p=.04$).

Therefore, with regards to dual-tasking in quiet and in noise, there appears to be subjective corroboration. However, the subjective ratings revealed a distinction between the two SNRs, which was not exhibited within the visual accuracy scores. This may be due to some insensitivity within the digit stream dual-task measure. Before this can be concluded, the auditory accuracy scores need to be assessed.

3.9.4.3. *Stability of primary task performance:* *Auditory accuracy*

When noise was introduced, all NH controls deteriorated in auditory accuracy, from originally being at ceiling (in quiet), with further decreases when noise was at -6dB SNR (Figure 3.12 a). On the other hand, at a given noise level, auditory accuracy did not appear to change substantially when participants were dual-tasking compared to performing just the auditory task. Indeed, the median difference in auditory accuracy between auditory-only and dual-task remained close to zero in both noise conditions (Figure 3.12b). However, it needs to be noted that there was considerable variability in how stable auditory performance was across participants (within each SNR).

It has been previously reported that, within the normal hearing population, there is a natural variability in the ability to discriminate speech in noise (Hällgren et al., 2001; Pichora-Fuller et al., 2016; Mattys et al., 2012). Therefore, it could be the case that the observed variation in auditory accuracy within the digit stream dual-task could simply be natural individual differences. This then introduces the concern that the NH controls will each respond differently to the acoustic challenge presented by the fixed noise levels (i.e. 0dB and -6dB SNR) across all NH controls (with some NH controls being unfairly disadvantaged). This, in turn, may cause inappropriate resource allocation.

This was tested by examining primary task stability via a 2x2 mixed effects logistic regression (which was applied to the auditory accuracy percentage). Both predictors were categorical: task type (i.e. auditory-only or dual-task) and noise type (i.e. 0dB or 6dB SNR). Interaction was allowed in this model, and participants were also set as random effects.

This model revealed an expected main effect of SNR: as SNR decreased, auditory accuracy significantly deteriorated ($F(1,108)=2251.1$, $p<.001$). The model also revealed a significant main effect of task, indicating that auditory accuracy significantly differed between auditory-only and dual-tasking conditions, although this effect was smaller and barely reached significance ($F(1,108)=4.1$, $p=.046$). Importantly, there was a significant interaction ($F(1,108) = 18.3$, $p<.001$). Therefore, post-hoc pairwise contrasts (Bonferroni-corrected) were pursued on the mean auditory accuracy in each condition (Table 3.6).

These revealed a significant improvement in auditory accuracy when dual-tasking at -6dB SNR ($p<.001$). However, at the easier noise level of 0dB SNR, auditory accuracy deteriorated when dual-tasking, although this change did not reach significance ($p=.08$). Thus, there appears to be instability in primary task performance. Consequently, the LE scores obtained are potentially unreliable.

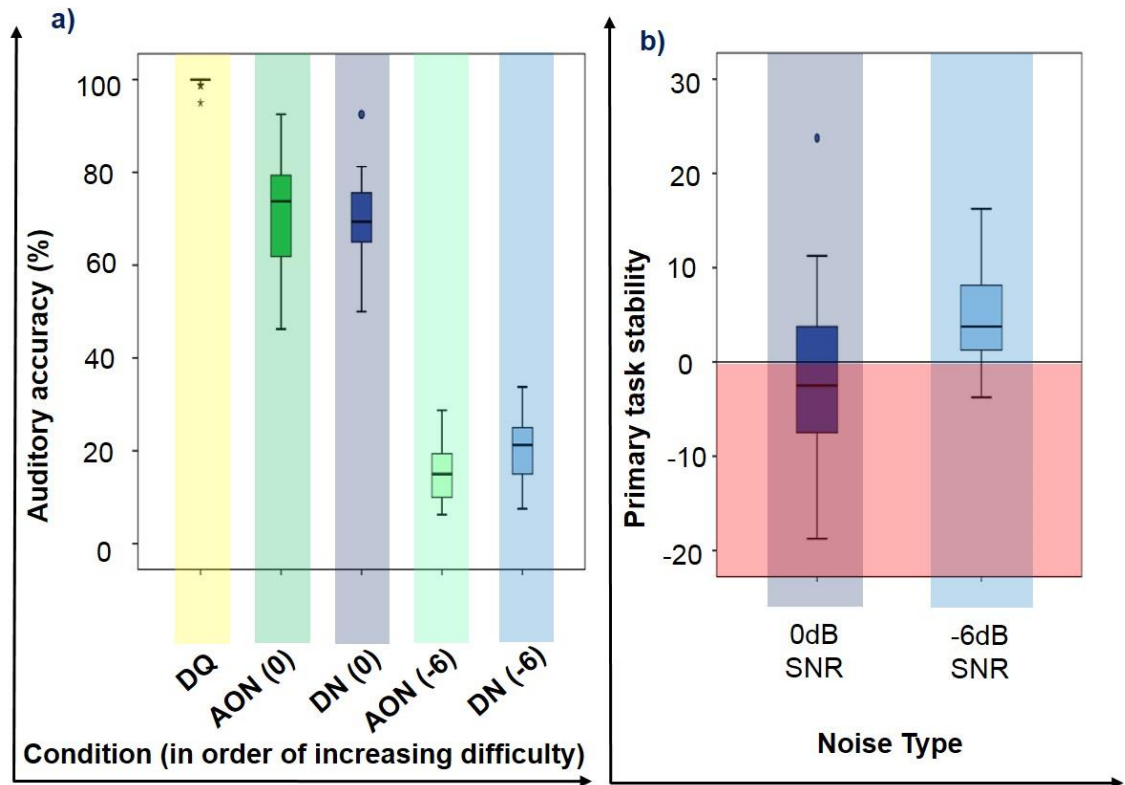


Figure 3.12: Boxplots of auditory accuracy (a) and primary task stability (b).

DQ = Dual-tasking in quiet, AON (0) = Auditory task only in noise of 0dB SNR, DN (0) = Dual-tasking in noise of 0dB SNR, AON (-6) = Auditory task only in noise of -6dB SNR, DN (-6) = Dual-tasking in noise of -6dB SNR. Red arrow indicates direction in which LE increases. Red shading either indicates deterioration in auditory accuracy when dual-tasking.

Unit of primary task stability = Difference in percentage.

Table 3.6: Mean and standard error of auditory accuracy for each test condition

Test condition	Mean	Standard error
Auditory only at 0dB SNR	71.9	2.1
Dual-tasking at 0dB SNR	69.6	1.8
Auditory only at -6dB SNR	15.3	1.0
Dual-tasking at -6dB SNR	20.0	1.3

3.9.4.4. Evaluation of viability of LE measurement

However, before the credibility of the LE scores can be definitively concluded, there is the need to confirm that the earlier methodological issue of over-reporting had not recurred. Thus, the nature of the NH controls' visual responses were more specifically interrogated. This was achieved by tallying up the number of trials for each possible type of visual response (i.e. accurate response, reduced target detection, and over-reporting) for each target presentation frequency (Figure 3.13).

It was revealed that the NH controls' responses shifted towards that of reduced target detection (i.e. targets were being missed) as more targets were presented in the digit stream. This was as expected. However, over-reporting still persisted within the participants' responses.

Therefore, the instances of over-reporting were isolated from the dataset for each test condition and target presentation frequency (Figure 3.14a). It was discovered that over-reporting occurred in the majority of the 28 participants and within all test conditions. Not only this, there appeared to be a tendency that, as test conditions became more difficult, incidence of over-reporting increased (Figure 3.14b). In addition, as the number of targets presented in the digit stream increased, over-reporting incidence also decreased (Figure 3.14c).

To better understand how the over-reporting phenomenon is affected by the test condition, and also the number of targets presented in a given digit stream, these two variables were set as predictors for a mixed effects logistic regression model. Test condition was treated as a categorical variable (i.e. visual-only, dual-tasking in quiet, dual-tasking at 0dB SNR and dual-tasking at -6dB SNR), whilst the number of targets was treated as continuous. Interaction was allowed in this model, and participants were also set as random effects.

This model revealed that over-reporting incidence differed significantly across the main effect of test condition ($F(3, 440)=8.2, p<.001$). In addition, the occurrence of over-reporting significantly decreased as the number of targets increased ($F(1,440)=94.4, p<.001$). Not only this, there was also a significant interaction ($F(3,440)=3.7, p=.012$). Therefore, it appears that the decrease in over-reporting with target number is different in the different test conditions. This may partly be due to the fact that over-reporting is relatively rare when there were 3 and 4 targets in the digit stream. Thus, change in over-reporting incidence across test condition is more likely to be detected when there were 1 or 2 targets present (compared to 3 or 4 targets). Since over-reporting is still occurring, and in a systematic manner, this adds to the concern that the LE scores may be unreliable.

3.9. Experiment 3

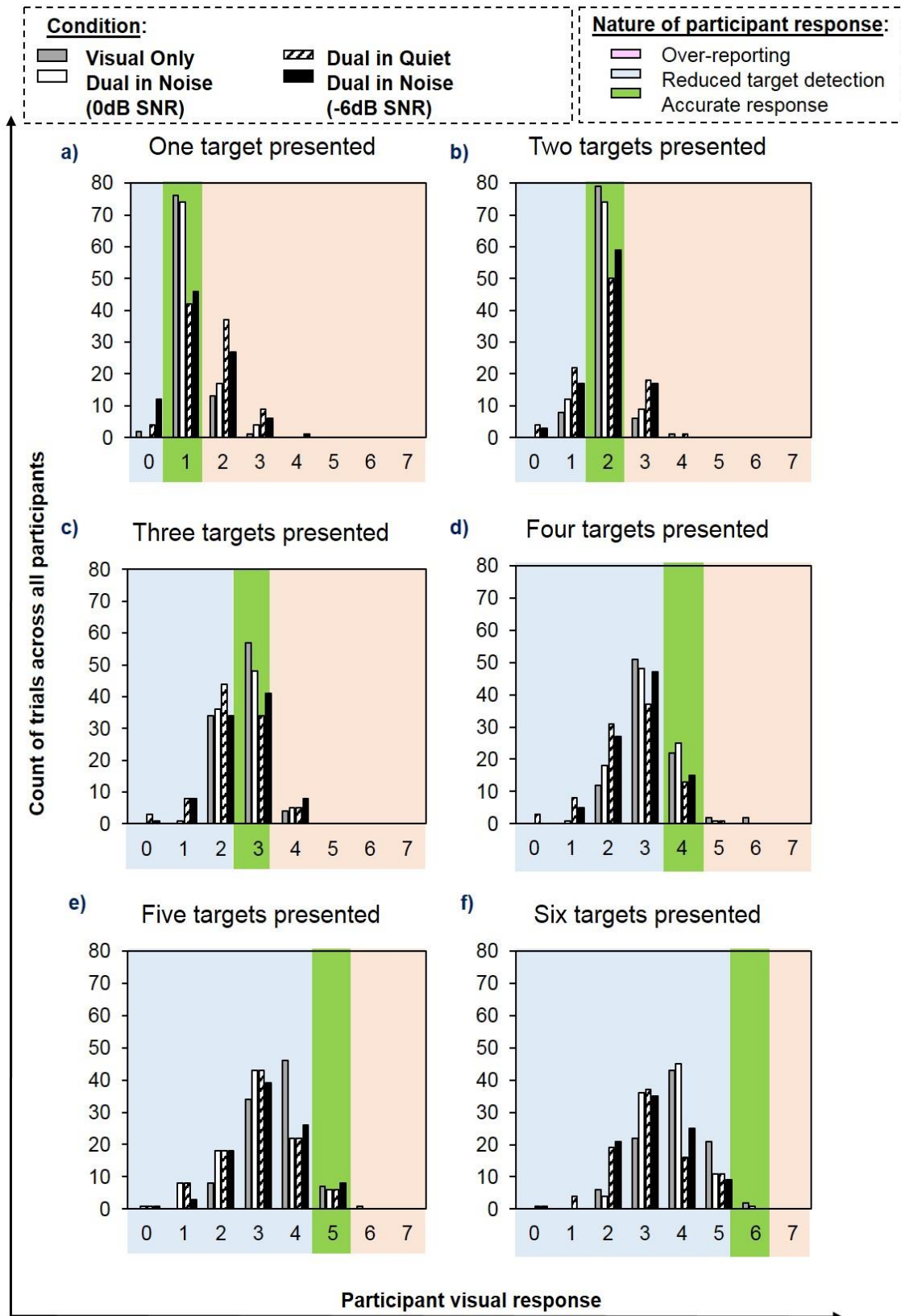


Figure 3.13: Histograms depicting the nature of participants' visual responses across trials (in each test condition) for each target presentation frequency.

3.9. Experiment 3

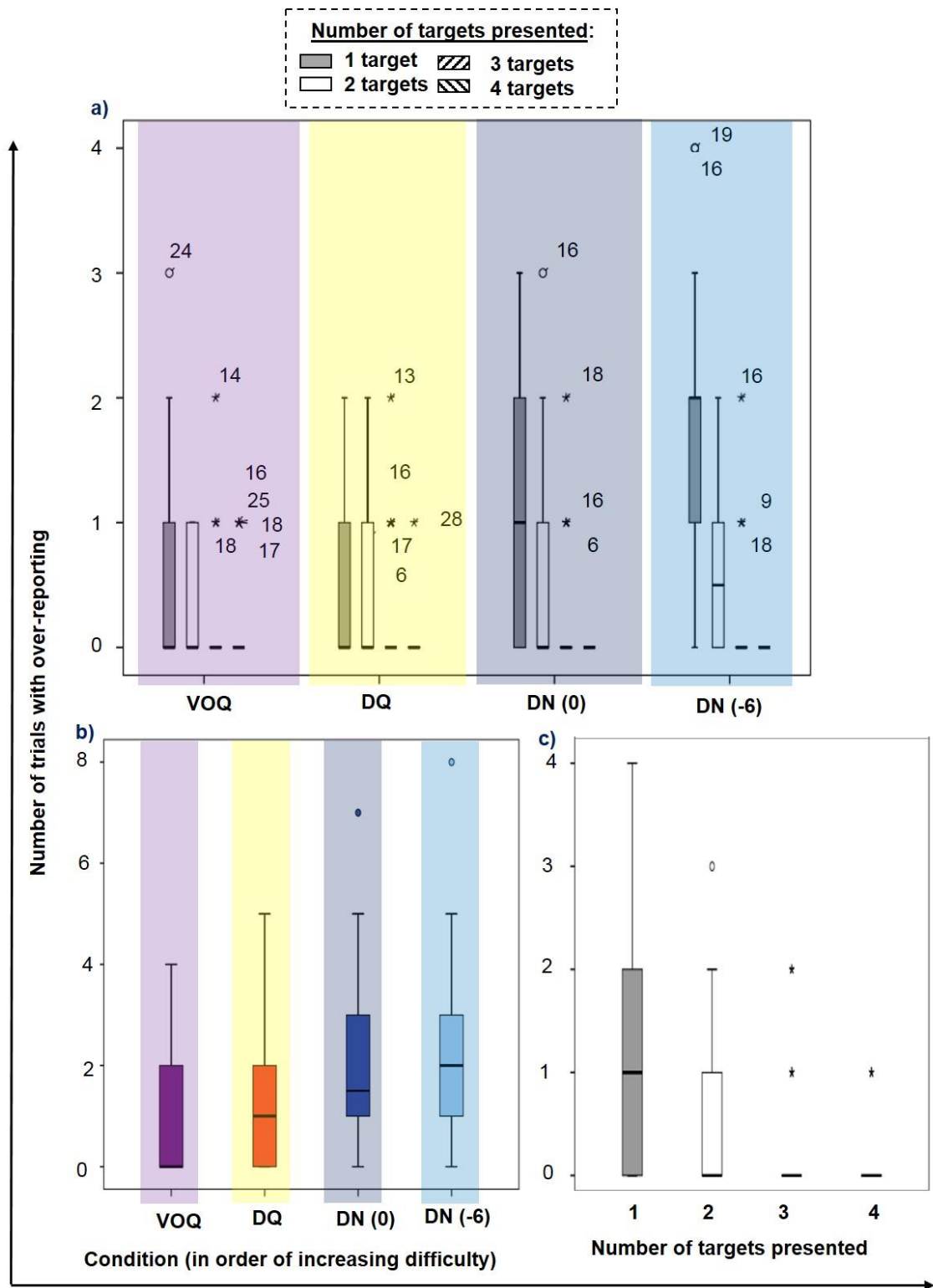


Figure 3.14: Boxplots of incidence of over-reporting, across all test conditions and target presentation frequency (**a**), for each test condition (summed across all conditions of target presentation frequency: **b**), and for each target presentation frequency (averaged across all test conditions: **c**). Total number of trials for each test condition is 18. VOQ = Visual task only in quiet, DQ = Dual-tasking in quiet, DN (0) = Dual-tasking in noise of 0dB SNR, DN (-6) = Dual-tasking in noise of -6dB SNR.

Asterisks (*) indicate that the occurrence of a particular value was an outlier. Superimposed numbers relate to the numerical identity of NH control(s) responsible for the given outlier value. Circles indicate extreme values.

3.10. Conclusions of Experiment 3

Some results in Experiment 3 were promising: visual accuracy was found to be significantly greater when performing the visual task alone, compared to dual-tasking in 0dB SNR and -6dB SNR. Visual accuracy was also significantly greater when dual-tasking in quiet compared to when dual-tasking in noise at -6dB SNR (thus producing positive LE scores).

Other results were cause for concern. For instance, over-reporting was still occurring. Also, the credibility of the LE scores has been marred by the presence of primary task instability. It was noticeable that there was variability in auditory performance across participants, which could be due to individual differences in auditory processing. This could, in turn, give rise to the possibility of there also being individual differences in cognitive capacity and, therefore, the ability to perform the dual-task. This could explain the primary task instability, as this is calculated as the difference between single and dual-task performance. This then has the implication that participants' performances can no longer be directly compared to one another. This is because some participants may be comfortably able to dual-task, whilst others are struggling.

Therefore, before the full dismissal of the digit stream dual-task for LE measurement, the next chapter will report an experiment that attempts to tailor the dual-task to accommodate these individual differences.

CHAPTER FOUR:

Introducing adaptive tracking into the refined digit stream dual-task (Experiment 4)

4.1. Introduction

As concluded in the previous chapter, the sensitivity of the digit stream dual-task appears to be particularly undermined by primary task instability. It was proposed that this instability in auditory performance could have arisen from a natural individual variation in auditory processing, which then impacts on cognitive capacity and, therefore, the ability to dual-task.

4.2. Proposed solution

To overcome any individual differences in neurocognition, there needs to be a tailoring of the digit stream dual-task for each participant. This tailoring does not necessarily require changes to the primary and secondary tasks themselves. After all, if the core of the issue is a variability in aptitude for speech-in-noise perception, then this can be controlled for by the signal-to-noise ratio (SNR) of the speech-shaped noise mask.

Accordingly, a potential approach to glean an appropriate SNR is to find each individual participant's speech reception threshold (SRT) and then base the SNRs of the speech-shaped noise mask on this value. This speech reception threshold can be defined as the lowest SNR at which speech is recognised, as denoted by the participant's speech intelligibility score (Plomp & Mimpen, 1979; Drullman et al., 1994). This ascertainment of SRT can be efficiently achieved by using the method of adaptive tracking (Leek, 2001; Wichmann & Hill, 2001; Levitt, 1971). An adaptive tracking procedure is any protocol whereby the stimulus level on any one trial is determined by the preceding stimulus and response (Levitt, 1971; Leek, 2001). Many well-known psychophysical techniques (such as the Methods of Limits) are essentially adaptive procedures, making this methodology a well-grounded approach to pursue (Levitt, 1971).

Because of the need to confirm the effectiveness of the adaptive tracking in controlling participant performance, the digit stream dual-task will not be trialled again for LE measurement yet. Instead, another larger-scale study will be executed (with a cohort of 30), whereby an adaptive tracking protocol will be integrated. The output SRTs will then be utilised to derive two bespoke noise levels for each participant.

To establish the efficacy of this attempt in tailoring the acoustic challenge, these bespoke noise levels will be incorporated within the dual-tasking test conditions to produce a refined digit stream paradigm. In order to simplify data interpretation at this early stage of development and LE test design, the homogeneity of the sample group has been maximised by again having exclusively normal hearing participants.

4.3. Experiment 4

4.3.1. Hypotheses

The adaptive tracking protocol will be successful in achieving the following:

- Obtaining a representative SRT (i.e. the SNR at which the participant has 50% auditory accuracy);
- Equalising the auditory performance level across the entire participant group.

When the participant is dual-tasking, visual accuracy will be affected according to their bespoke SNR such that:

- Visual accuracy will decrease when the SNR decreases.

This change in visual accuracy will not be at the cost of resource allocation:

- Auditory accuracy will remain constant when the participant is dual-tasking (relative to that achieved when performing only the auditory task).
- Over-reporting will also be ameliorated.

There will also be subjective corroboration (i.e. that of perceived effort) of the behavioural performance:

- As SNR decreases (and visual accuracy decreases), subjective ratings will become higher.

4.3.2. Methods

4.3.2.1. Participants

Thirty adult normal hearing (NH) controls (19 females and 11 males, mean age: 45.93 years, \pm 15.17 S.D) were recruited. All participants had normal, or corrected-to-normal vision, and were native British English speakers.

4.3.2.2. Procedure

Location of testing: Testing sessions were conducted either within the participants' homes or offices.

Adaptive tracking: Each SRT test consisted of 20 randomly selected BEL sentences (i.e. the same sentence corpus utilised for the Digit Stream Dual-Task). As the SRT was defined to be the SNR at which the speech intelligibility score becomes 50% correct, this translated to the participant being able to understand (and type) 50% of the presented key words (note, each sentence had 4 key words).

The starting SNR was set at 0dB, and adjusted up or down by 10dB before the first reversal (i.e. when the participant's accuracy in their answer opposed that given in the previous trial) and 6.5dB before the second reversal. From the second reversal onwards, the step size was 3dB (note, the increments or decrements by which the stimulus is increased, or decreased, are referred to as steps: Levitt, 1971).

The selection of SNR (i.e. the adaptive algorithm) was such that, if more than 2 key words were correct (out of the total of 4), the SNR decreased on the next trial; if less than 2 key words were correct, the SNR increased; and if 2 key words were correct, the SNR remained the same. The SRT was calculated from the mean of all levels visited (in SNR) at the final even number of reversals with the 3dB step size.

The participant's SRT was then used to set the SNR for the first of the two noise levels tested within the Digit Stream Dual-Task. For the second noise level, the applied SNR was set to be 4dB above the participant's SRT (in order to create a marginally easier noise level for comparison purposes).

MATLAB (R2015b, MathWorks) was utilised to deliver these BEL sentences, implement the adaptive algorithm, and also collect the participants' responses (via a similar GUI as used for the auditory response within the Digit Stream Dual-Task). Once a particular BEL sentence was used, it was not delivered again for the remainder of the testing session.

Psychometric functions were also produced for each NH control, using R (R-3.4.0, The R Foundation). A typical psychometric function relates a participant's performance to an independent variable (which is usually the physical quantity of a stimulus within the context of psychophysical research: Wichmann & Hill, 2001; Klein, 2001; Levitt, 1971; Leek, 2001). In this instance, the psychometric abscissa was the SNR and the ordinate represented the observer's response. A sigmoid curve was expected to comprise the shape of this psychometric function (Levitt, 1971).

Acoustic challenge: The same as for Experiment 3 (Chapter 3), but using the individually adapted SNRs.

Primary task of the Digit Stream Dual-Task: The same as for Chapter 3.

Secondary task of the Digit Stream Dual-Task: The same as for Chapter 3, except with the following modifications: the target was now the digit “8” and only the digits 1, 2, 4, 5, 6, 7, 8 and 9 could be presented within the digit stream. This was part of the attempt to strengthen the secondary task against the contamination of over-reporting. It was hoped that the digit “8” was a relatively salient stimulus (with its symmetry in the horizontal and vertical planes), which then might assist with detection (especially since the digits “0” and “3” had been removed from the possible stimulus options due to the similarity of their written form with the target “8”).

Other LE indices: The same as for Experiment 3 (i.e. 5-point Likert scale for subjective ratings).

Test conditions (and trial structure): In total, there were 4 possible test conditions for the Digit Stream Dual-Task (Figure 4.1), of which there was one condition for speech type (i.e. normal), as well as two conditions for task type (i.e. auditory task only, or dual-task) and two conditions for noise type (i.e. SNRs at the SRT and at 4dB above SRT). There were 18 trials for each test condition. The trial structure remained consistent with that used for Experiment 3.

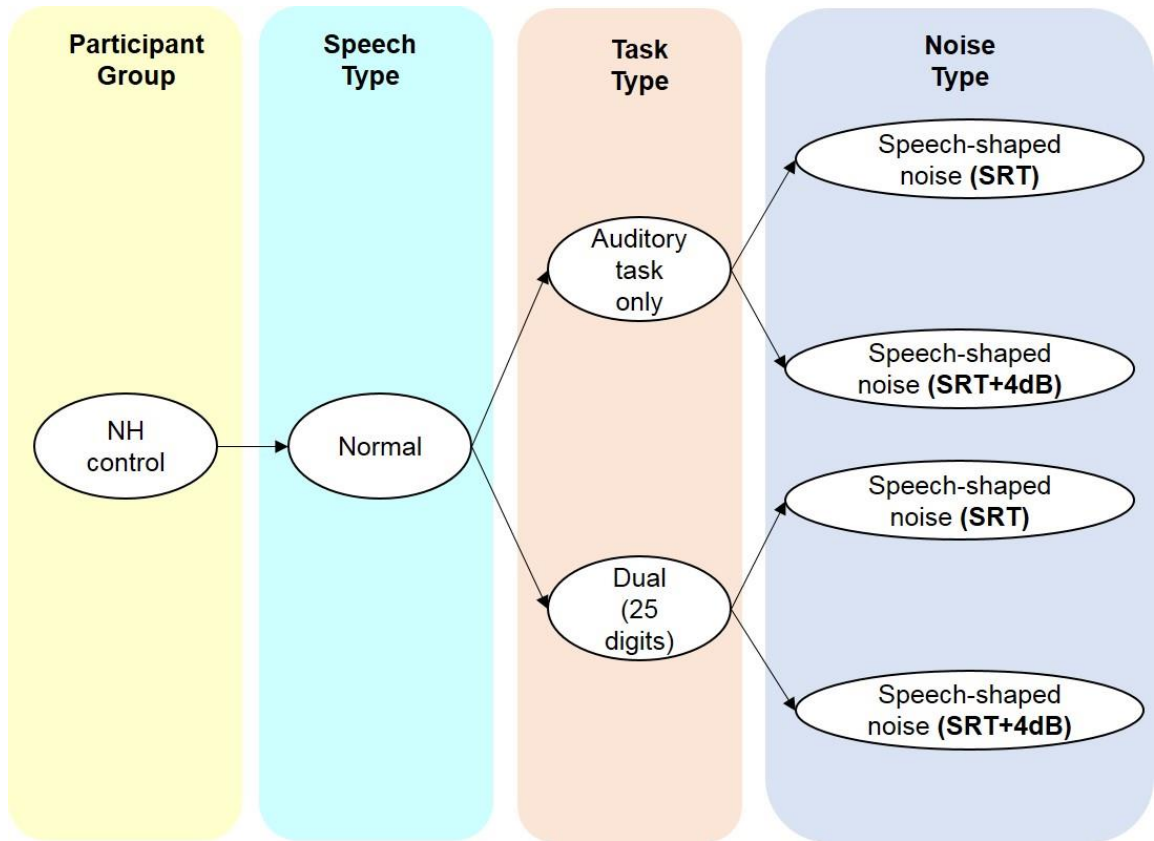


Figure 4.1: A schematic of test conditions (same structure as for Figure 3.1).

Protocol: For the overall protocol of the testing session, see Table 4.1. The participants first had the opportunity to familiarise themselves with the Digit Stream Dual-Task, via a practice run of the single-task (i.e. the primary and secondary tasks on their own, 5 trials each) and dual-task condition (both in quiet and in noise of 0dB SNR, 5 trials each). Adaptive tracking was then executed, in order to ascertain the bespoke SNRs for each participant before the test run was pursued. During the test run itself, the participant had a warm-up before each test condition (consisting of 5 trials).

The order of the test conditions were randomised and counter-balanced for each participant (i.e. steps 6-9 in Table 4.1), as also were the lists of sentence stimuli (to ensure no repetition).

4.3. Experiment 4

Table 4.1: Protocol for Experiment 4. Same detail provided as for Table 2.3, including red numbers (with asterisks) indicating all the steps in the protocol where the order of the test conditions were randomised and counter-balanced together.

Step	Training, Practice or Test?	Task Type	Speech Type	Noise Type	Total number of trials
1	<i>Practice</i>	Visual task only at 25 digits	N/A	Quiet	5 trials
2	<i>Practice</i>	Auditory task only	Normal	Quiet	5 trials
3	<i>Practice</i>	Dual at 25 digits	Normal	Quiet	5 trials
4	<i>Practice</i>	Dual at 25 digits	Normal	Speech-shaped noise at 0dB SNR	5 trials
5	Test	Adaptive tracking (tracking 50% correct)	Normal	Speech-shaped noise at variable SNR	20 trials
6 *	<i>Practice</i>	Auditory task only	Normal	Speech-shaped noise at SRT	5 trials
	Test	Auditory task only	Normal	Speech-shaped noise at SRT	20 trials
7 *	<i>Practice</i>	Auditory task only	Normal	Speech-shaped noise at SRT +4dB	5 trials
	Test	Auditory task only	Normal	Speech-shaped noise at SRT +4dB	20 trials
8 *	<i>Practice</i>	Dual at 25 digits	Normal	Speech-shaped noise at SRT	5 trials
	Test	Dual at 25 digits	Normal	Speech-shaped noise at SRT	18 trials
9 *	<i>Practice</i>	Dual at 25 digits	Normal	Speech-shaped noise at SRT +4dB	5 trials
	Test	Dual at 25 digits	Normal	Speech-shaped noise at SRT +4dB	18 trials

Statistical analysis: The same as for Experiment 3 (Chapter 3), with the addition of the Pearson test of correlation.

4.3.3. Results and Discussion: Descriptive and inferential statistics

4.3.3.1. Secondary task performance: Visual accuracy

It was found that visual accuracy deteriorated when the SNR worsened (Figure 4.2). This was as hypothesised. To ascertain the statistical significance of this deterioration, a mixed effects logistic model with a single predictor of test condition (i.e. dual-tasking in the two noise levels) was applied to target detection percentage. Participants were set as random effects, as part of the attempt to generalise the findings to the population. A statistically significant result ($F(1,58)=41.9, p<.001$) suggests that the visual accuracy did differ across SNR. Because the change in visual accuracy between two subtly different noise levels was above chance, it thus seems probable that statistically significant differences would also be detected if LE was calculated (i.e. when visual accuracy at each noise level was compared to that achieved when dual-tasking in quiet).

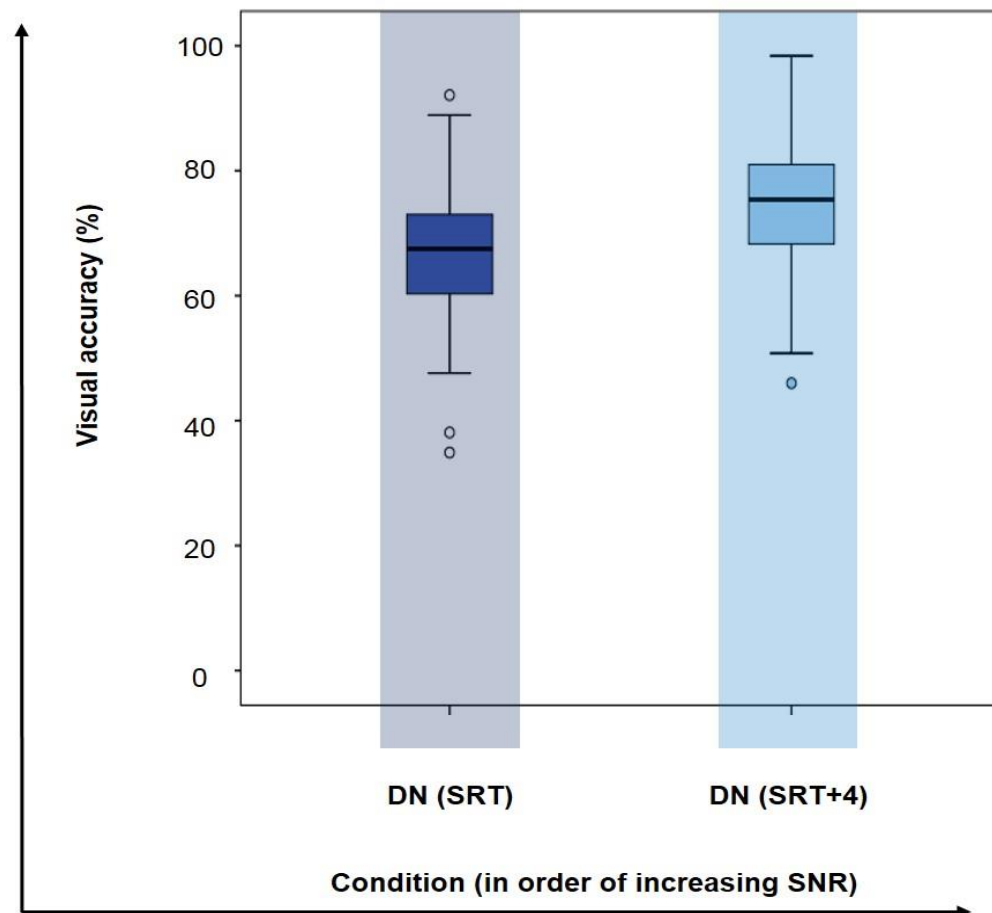


Figure 4.2: Boxplots of visual accuracy across test conditions (with target detection percentage corrected for any instances of over-reporting). DN (SRT) = Dual-tasking in noise at SRT, DN (SRT+4) = Dual-tasking in noise at 4dB above SRT.

4.3. Experiment 4

4.3.3.2. Evaluation of validity of secondary task performance: Subjective ratings

The majority of participants exhibited an increase in subjective ratings (i.e. towards the “extremely hard work” end of the rating scale) as the noise level increased (Figure 4.3). This data trend was as hypothesised, i.e. there appears to be subjective corroboration.

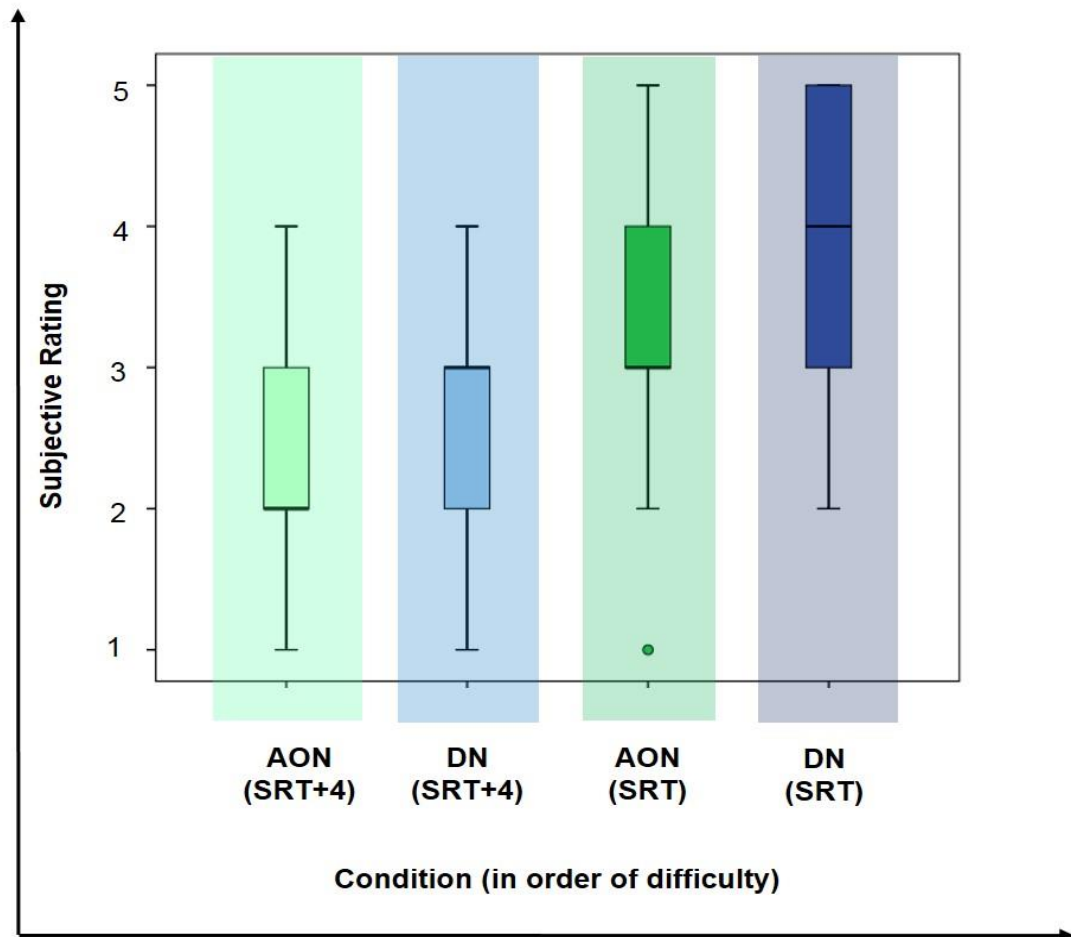


Figure 4.3: Boxplots of subjective ratings across test conditions.

AON (SRT) = Auditory task only in noise at SRT, DN (SRT) = Dual-tasking in noise at SRT, AON (SRT+4) = Auditory task only in noise at 4dB above SRT, DN (SRT+4) = Dual-tasking in noise in noise at 4dB above SRT.

Subjective ratings: 1= Not at all hard work, 2 = Quite hard work, 3 = Medium hard work, 4 = Very hard work, 5 = Extremely hard work.

To establish the statistical significance of this trend, a mixed effects linear model was applied, with a single predictor of test conditions (with four levels). Participants were set as random effects. A statistically significant result ($F(3,87)=33.9$, $p<.001$) suggests that subjective ratings differed across test condition. Therefore, post-hoc pairwise comparisons (Bonferroni-corrected) were pursued on the mean subjective rating for each condition (Table 4.2).

Table 4.2: Mean and standard error of subjective ratings for each test condition

Test condition	Mean	Standard error
Auditory only at SRT+4dB	2.1	0.1
Auditory only at SRT	3.1	0.2
Dual-tasking at SRT+4dB	2.8	0.2
Dual-tasking at SRT	3.9	0.2

All but one of these comparisons yielded significant differences ($p=.003$ when dual-tasking at 4dB above SRT was compared to auditory-only at the same SNR, $p<.001$ for remainder of comparisons). The exception was when comparing the auditory-only condition with noise set at the SRT to the dual-tasking condition when noise was 4dB above SRT ($p=.36$). Thus, as test condition became progressively harder, subjective ratings significantly increased until the SNR was at its most difficult (i.e. the SRT).

4.3.3.3. *Stability of primary task performance:* *Auditory accuracy*

Auditory accuracy scores (Figure 4.4 a) revealed that the majority of NH controls exhibited enhanced performance when the SNR improved by 4dB relative to their SRT (for the individual SRT scores as obtained by the adaptive tracking, see Figure 4.4 b).

To ascertain whether or not primary task stability has been achieved (Figure 4.4c), a 2x2 mixed effects logistic regression was applied to auditory accuracy percentage. Both predictors were categorical: task type (i.e. auditory-only or dual-task) and noise type (i.e. at SRT, or 4dB above SRT). Interaction was allowed in this model, and participants were also set as random effects.

This model revealed an expected main effect of SNR: as SNR increased, auditory accuracy significantly improved ($F(1,116)=1216.9$, $p<.001$). The main effect of task, however, was not significant ($F(1,116) = 1.1$, $p=.29$), indicating a lack of difference between auditory-only and dual-tasking conditions. There was also no significant interaction ($F(1,116)=2.3$, $p=.13$).

This lack of task effect is desirable, as this may point towards stability in primary task performance. However, it does need to be noted that a statistically null result does not equate to evidence of absence. Alternative approaches to demonstrate primary task stability are to use equivalence testing (such as the two one-sided test procedure, i.e. TOST), and/or measures of agreement (e.g. the Bland and Altman plot). However, these techniques are difficult to implement within this experiment, because there is a requirement to define (in advance) what an acceptable level of difference is. In this case, the level of difference applies to the point at which primary task stability becomes instability. There is not enough knowledge, or prior data, to set this criterion (of the point of instability) for this experiment with sufficient confidence.

4.3. Experiment 4

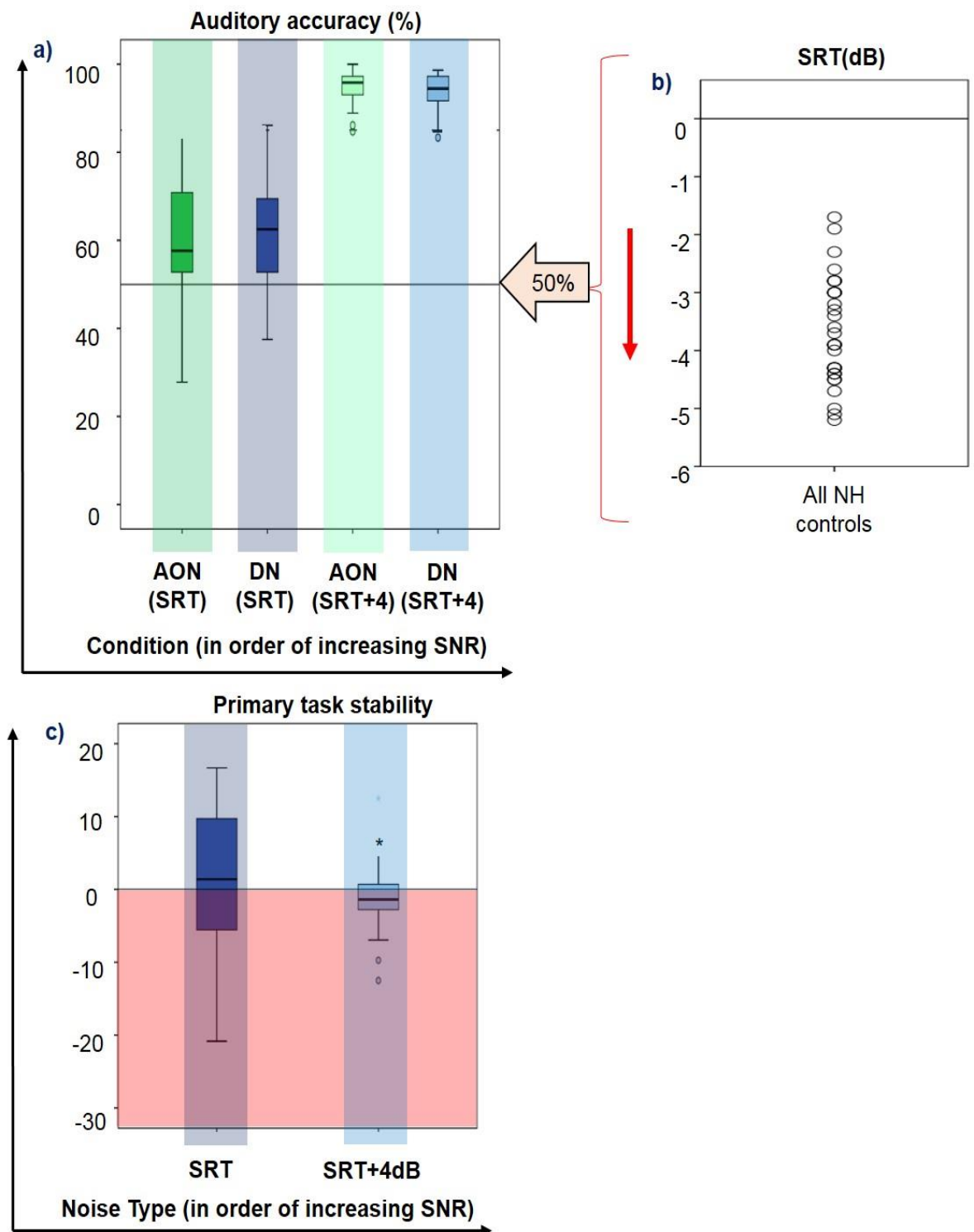


Figure 4.4: Boxplots of auditory accuracy **(a)** and primary task stability **(c)**, and dot plot of obtained SRTs **(b)**.

AON (SRT) = Auditory task only in noise at SRT, DN (SRT) = Dual-tasking in noise at SRT, AON (SRT+4) = Auditory task only in noise at 4dB above SRT, DN (SRT+4) = Dual-tasking in noise 4dB above SRT.

Red shading indicates deterioration in auditory performance when dual-tasking.
Red arrow indicates superiority of SRT.

Unit of primary task stability = Difference in percentage.

4.3. Experiment 4

4.3.3.4. Evaluation of potential viability of Digit Stream Dual-Task as LE measurement

With detection of statistically significant differences in secondary task performance between two subtly different noise levels, as well as apparent stability in primary task performance, the tailored Digit Stream Dual-Task appears to show promise as a viable LE measure. However, before this could safely be concluded, there was a need to check whether the visual accuracy scores had yet again been compromised by over-reporting (i.e. the participant reporting more targets than actually were presented). Similar to Experiment 3, the number of trials for each type of visual response were tallied up for each target presentation frequency in each test condition (i.e. accurate response, reduced target detection, or over-reporting).

It emerged that, as the number of targets presented in the digit stream increased, the participants' visual response tended to drift towards that of reduced target detection (Figure 4.5). This was as expected. Unfortunately, it was also revealed that over-reporting was still occurring.

Therefore, the instances of over-reporting were isolated from the dataset for each test condition and target presentation frequency (Figure 4.6a). It was discovered that over-reporting occurred in the majority of the participants and within all test conditions. Not only this, there appeared to be the tendency that, as the SNR increased, the incidence of over-reporting decreased (Figure 4.6b). In addition, as the number of targets presented in the digit stream increased, over-reporting incidence also decreased (4.6c). To ascertain the statistical significance of the effect of test condition and target presentation frequency on over-reporting, these variables were set as predictors for a mixed effects logistic regression model. Test condition was treated as a categorical variable (i.e. dual-tasking at the two different noise levels), whilst the number of targets presented in the digit stream was treated as continuous. Interaction was allowed, and participants were set as random effects.

This model revealed a significant main effect of test condition ($F(1,176)=7.1, p=.01$), as well as a significant main effect of target presentation frequency ($F(1,176)=32.5, p<.001$). Thus, over-reporting significantly decreased when SNR improved, and also when the number of targets increased. In addition, there was a significant interaction ($F(1,176)=4.5, p=.04$), indicating that the decrease in over-reporting with target number was different for each SNR. This may partly be due to the fact that over-reporting is relatively rare when there were 2 and 3 targets in the digit stream, compared to when there was only 1 target. Therefore, change in over-reporting incidence across SNR is more likely to be detected when there was only one target in the digit stream.

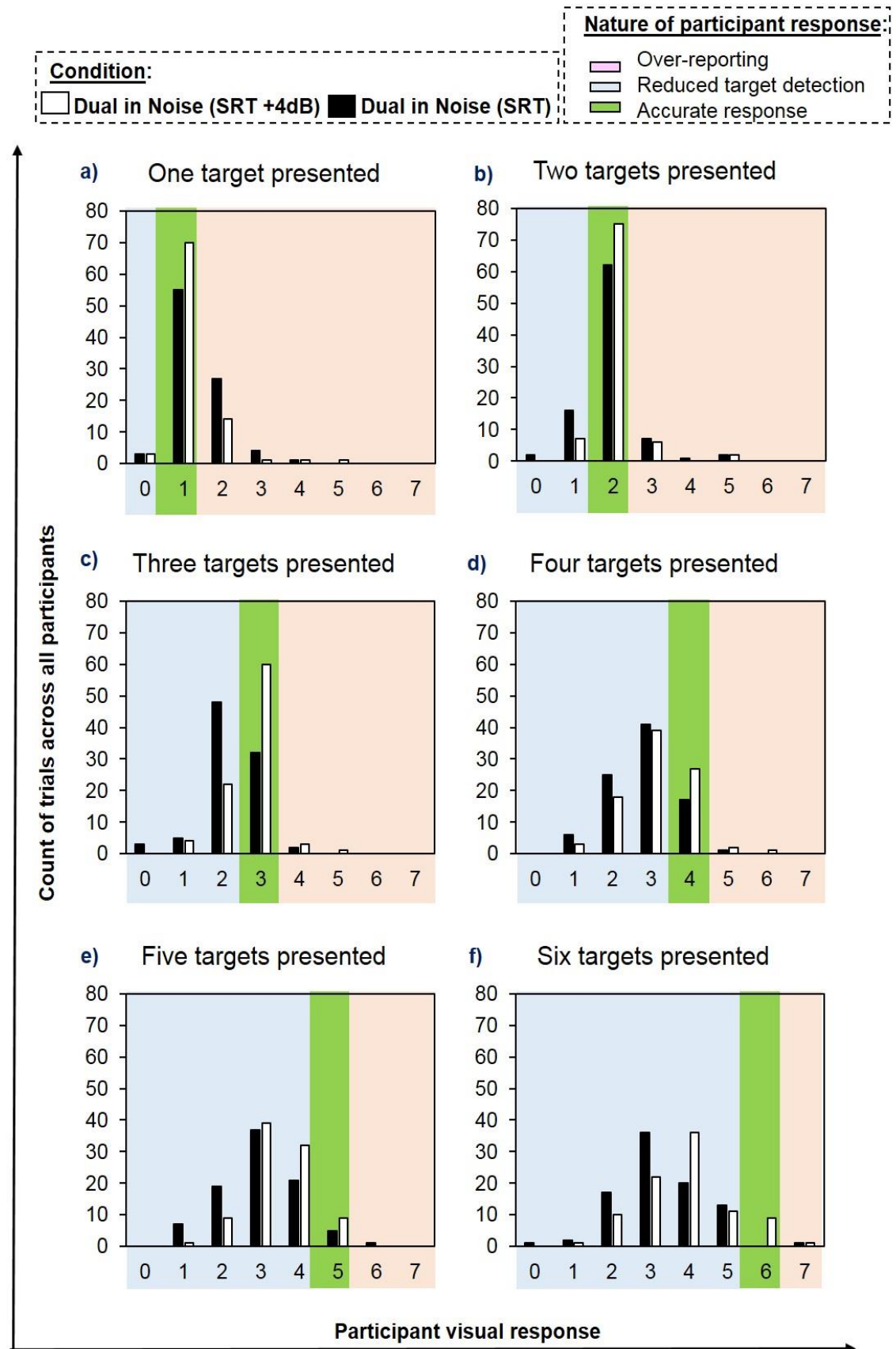


Figure 4.5: Histograms depicting the nature of participants' visual responses across trials (in each test condition) for each target presentation frequency.

4.3. Experiment 4

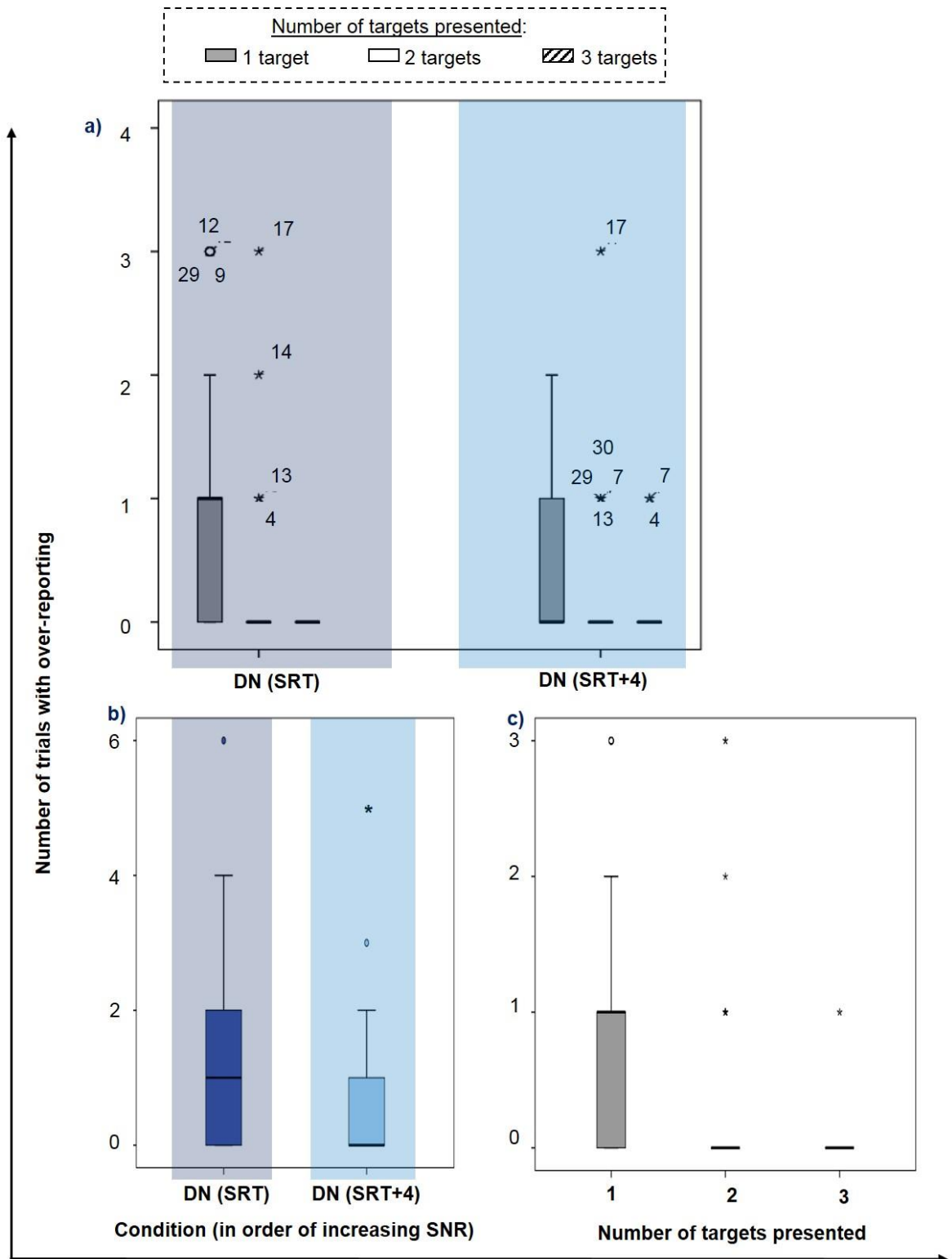


Figure 4.6: Boxplots of incidence of over-reporting, across all test conditions and target presentation frequency (a), for each test condition (summed across all conditions of target presentation frequency: b), and for each target presentation frequency (averaged across all test conditions: c). Total number of trials for each test condition is 18.

DN (SRT) = Dual-tasking in noise at SRT, DN (SRT+4) = Dual-tasking in noise at 4dB above SRT.

Asterisks (*) indicate that the occurrence of a particular value was an outlier. Superimposed numbers relate to the numerical identity of NH control(s) responsible for the given outlier value. Circles indicate extreme values.

Another aspect of the data that needed to be revisited was the auditory tracking data. After all, whilst there appears to be primary task stability, it is still necessary to confirm whether the tracking procedure had actually succeeded in its objective of equalising auditory performance level.

When the participant's SRT was plotted against the actual performance attained when performing only the auditory task within the Digit Stream Dual-Task (with noise set at the SRT – see Figure 4.7 a), it was found that less than a fifth of the NH controls were actually at the desired level of performance (i.e. 50% correct). Instead, there was considerable variability in auditory accuracy, ranging from 25% correct to even 85% correct. Not only this, it was also found that auditory accuracy had a significant positive correlation with the SNR applied (Pearson's $r(30)=0.733$, $p<.001$). This correlation should not have occurred, as the purpose of the adaptive tracking was to find SNRs that would fix the participant group's performance level at around 50%.

This inconsistency in auditory performance across the NH controls was in spite of reasonable tracks, and also the expected sigmoidal shape to psychometric functions (for an exemplar of psychometric function and affiliated track produced by NH control 12, see Figure 4.7 b and c respectively; for remainder of NH controls' psychometric functions and associated tracks, please see Appendices 4.1-4.8).

Therefore, there is now concern that the tracking had failed in producing an appropriate SNR to equalise the participants' auditory performance levels, with the majority of participants performing either worse or better than the intended performance level. The better auditory performances were equally worrying as the poorer performances, because it becomes difficult to determine whether the improvement in performance was due to additional LE being utilised (to cope with the acoustic challenge), or whether the SNR was actually too easy (and, therefore, was an inappropriate level of acoustic challenge).

However, there are other factors that need to be considered with regards to the higher auditory performance levels (relative to that of tracking). Notably, the SNR used in the dual-task was constant (compared to the continuously changing SNR used during adaptive tracking) which gives rise to the possibility that the participants may have adapted to the chosen acoustic challenge. This could then account for improved performance relative to that tracked.

However, it does need to be noted that this experiment involved the exclusive testing of normal hearing participants. This was a deliberate strategy, in an attempt to avoid any confusion in data interpretation that might be incurred by testing CI users (because of the high heterogeneity typical of the CI user population: Blamey et al., 2013; Green et al., 2007; Holden et al., 2013). Nonetheless, using NH controls does not guarantee that any observed behaviour would actually be representative of CI users themselves, especially when considering the limitations unique to electric hearing produced by the CI neuroprosthesis (e.g. Faulkner & Pisoni, 2013).

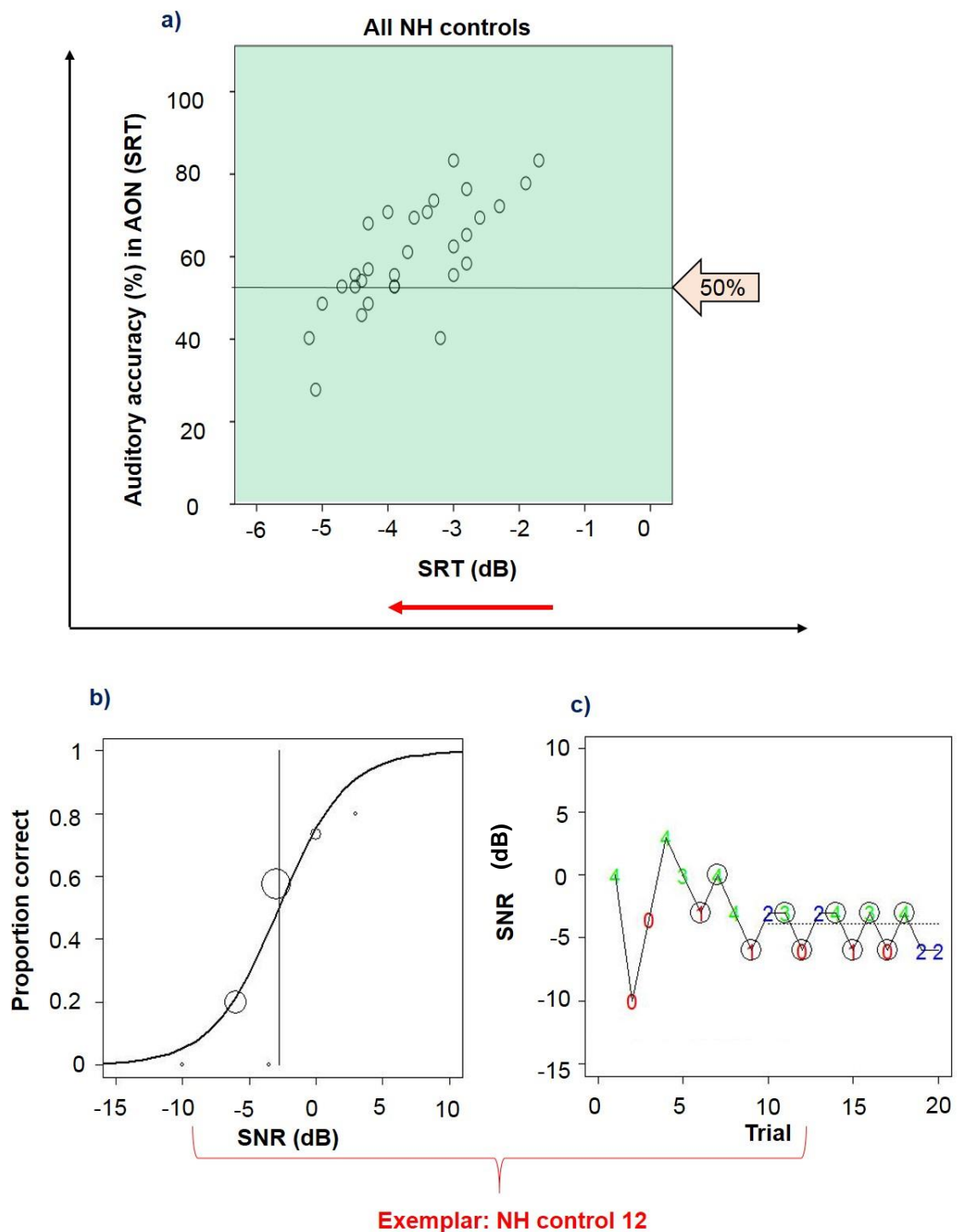


Figure 4.7: Scatterplot of auditory accuracy versus SRT for all NH controls **(a)**, and exemplars of a psychometric function and adaptive track **(b, c)**. AON (SRT) = Auditory only with noise with SNR at SRT. Red arrow indicates direction of superiority in SRT value **(a)**.

The psychometric function depicts the proportion of average number of key words of a BEL sentence the participant successfully heard (and typed) at a given SNR **(b)**. The adaptive track depicts how the SNR was changed across the trials according to the adaptive algorithm applied **(c)**.

Superimposed circles on psychometric function = Participant's response data, with size of circle controlled by the number of trials run at that particular SNR. Superimposed circles on adaptive track = Reversals. Superimposed numbers on adaptive track = Total number of key words correct, with colour indicating whether the track remained at the same SNR (blue), increased (red), or decreased (green) on the next trial.

4.4. Conclusions of Experiment 4

Despite this digit stream dual-task being able to detect statistically significant differences within visual performance in response to only a few decibels change in SNR (with no apparent cost to primary task stability), over-reporting continues to interfere with the visual response. In addition, it also seems to be the case that the adaptive tracking protocol had failed to equalise the auditory performance level in all participants. Therefore, the viability of this refined digit stream dual-task as a form of LE measurement is doubtful.

However, it was also highlighted that this experiment involved participants that were not necessarily entirely representative of the CI user population. Thus, it is not yet appropriate to reject the digit stream dual-task (or the tailoring of it) as a framework for a LE test until listening conditions become more relevant to the cochlear implant case. Accordingly, this adaptation of listening conditions will be pursued in the next chapter.

PART FOUR:

DIGIT STREAM

DUAL-TASK

(Cochlear Implant Simulations
& Cochlear Implant Users)

CHAPTER FIVE:

Modifying and using the refined digit stream dual-task in cochlear implant simulations testing (Experiments 5, 6 & 7)

5.1. Introduction

As concluded in the previous chapter (Chapter 4), the digit stream dual-task is still of questionable viability due to the continued presence of over-reporting, as well as a potential failure in adaptive tracking. However, it was also highlighted that these issues were being encountered in normal hearing participants in relatively normal (albeit noisy) conditions. Thus, the testing conditions used were not directly reflecting what occurs during electric hearing produced by the CI neuroprosthesis.

5.2. Proposed solution

It cannot be ignored that the original reason for exclusively testing normal hearing participants was to avoid the complication of participant heterogeneity that the CI user population is reputed for (Blamey et al., 2013; Green et al., 2007; Holden et al., 2013). Therefore, a compromise to create more pertinent listening conditions (i.e. related to the CI case) is to utilise simulations of the CI sound. One technique of generating these CI simulations is noise-excited vocoding (Faulkner et al., 2000; Loizou et al., 1999; Shannon et al., 1995). This technique mimics the bandpass filtering that is typically part of the CI processing and then multiplies the extracted temporal envelopes of each channel with an independent white noise matched to the bandwidth of that channel (Dorman et al., 1997; Rosen et al., 1999; Fu & Shannon, 1999; Shannon et al., 1998; 1999).

Via this technique, the spectral resolution of any sound can be explicitly manipulated, especially by controlling the number of bandpass filters permitted in the vocoder (Winn et al., 2015). Spectral resolution can be defined as the ability of the listener to distinguish sounds which differ in their spectral shape (Winn et al., 2015). For example, within a speech-related context, spectral resolution underlies the capacity to distinguish acoustically similar consonant pairs, such as /b/-/d/ or /t/-/k/ (Dubno et al., 1982; Munson et al., 2003).

The application of this technique has been found to induce levels of performance typical of CI users in people with normal hearing (Winn et al., 2015; Fu & Galvin, 2008; Rosen et al., 1999; Shannon et al., 1995).

Indeed, the impact of varying number and also spacing of the frequency bands (within the vocoding) on speech intelligibility for normal hearing listeners has been found to resemble the number and placement of electrodes in actual CI users (Friesen et al., 2001; Fu & Galvin, 2008; Shannon et al., 1995).

In order to accommodate the anticipated variability of speech-in-noise ability, the adaptive tracking protocol will continue to be applied in order to determine bespoke SNRs (Hällgren et al., 2001). There will be the addition of another protocol, but this time to tailor the level of spectral resolution to each participant. This is because there is also a similar variability in the ability of the listener to process degraded stimuli, such as noise-excited vocoded sound (Mattys et al., 2012; Peelle, 2017). This differing ability to process vocoded sound is potentially another factor which may influence the participant's cognitive capacity (and, therefore, their ability to dual-task). This, in turn, may cause an unfair disadvantage for some participants (in their auditory performance), especially for individuals who may struggle with the additional distortion of vocoded sound.

This new tracking will be conducted first (i.e. before the noise tracking), within quiet listening conditions, in order to obtain an appropriate number of “channels”, i.e. the number of frequency bands the auditory stimulus is divided into (the fewer the channels, the poorer the spectral resolution). The noise tracking itself will then attempt to find an appropriate SNR for the vocoded stimuli (that has already incorporated the bespoke number of channels for that specific participant). The intention is that the combination of these two types of tracking will not only control the acoustic challenge presented by the noise, but also the acoustic challenge presented by the vocoded stimuli themselves. Because the spectral degradation encountered within the vocoded stimuli is atypical for a naïve normal hearing listener, training will also be provided to the participants prior to any tracking and testing.

Since there are now new protocols with regards to the tracking, and also the training, there is the need to collect baseline data again. Accordingly, there will be two small-scale studies (i.e. Experiments 5 and 6) in order to establish all the necessary testing parameters (and trial them within the digit stream dual-task) before a larger-scale study is pursued (i.e. Experiment 7).

5.3. Experiment 5

5.3.1. Hypotheses

The use of the two adaptive tracking procedures will equalise the auditory performance level in quiet and in noise. When noise is introduced into the digit stream dual-task, the following will occur:

- Visual accuracy will decrease when the SNR decreases (i.e. LE will increase).

This change in visual accuracy will not be at the cost of resource allocation:

- Auditory accuracy will remain constant when the participant is dual-tasking (relative to that achieved when performing only the auditory task).
- Over-reporting will also be ameliorated.

In addition, there will be subjective corroboration of the behavioural data:

- Perceived effort will increase as LE scores increase.

5.3.2. Methods

5.3.2.1. Participants

Six adult normal hearing (NH) controls (4 females and 2 males, mean age: 46.83 years, \pm 13.61 S.D) were recruited. All participants had normal, or corrected-to-normal vision, and were native British English speakers.

5.3.2.2. Procedure

Location of testing: Testing sessions took place within the participants' home or office environments.

Primary task of the Digit Stream Dual-Task: As for Experiment 4.

Secondary task of the Digit Stream Dual-Task: As for Experiment 4.

Acoustic challenge: Same as before.

Vocoding parameters: The same procedure and software used in the Pilot Study was applied to implement the vocoding needed for all auditory stimuli within the two types of tracking, CI simulation training and also the Digit Stream Dual-Task. The only modification was that the number of analysis filters applied was determined by the channel number required by the task, or participant, at hand.

Other LE indices: The same subjective ratings as before.

CI simulation training: The same training software (and stimuli) from the Pilot Study were used. Training amounted to 15 minutes in total and comprised of three stages (Figure 5.1): 5 minutes in quiet, with 12 “channels” (i.e. frequency bands) within the vocoded stimuli; 5 minutes, still in quiet, but this time, the vocoding was set at the participant’s tracked number of channels (for details of the adaptive tracking, see the next section); and another 5 minutes, but this time in noise with the vocoding still set at the participant’s bespoke number of channels (the SNR being set at the participant’s bespoke level: for details of the adaptive tracking used for noise, see the next section).

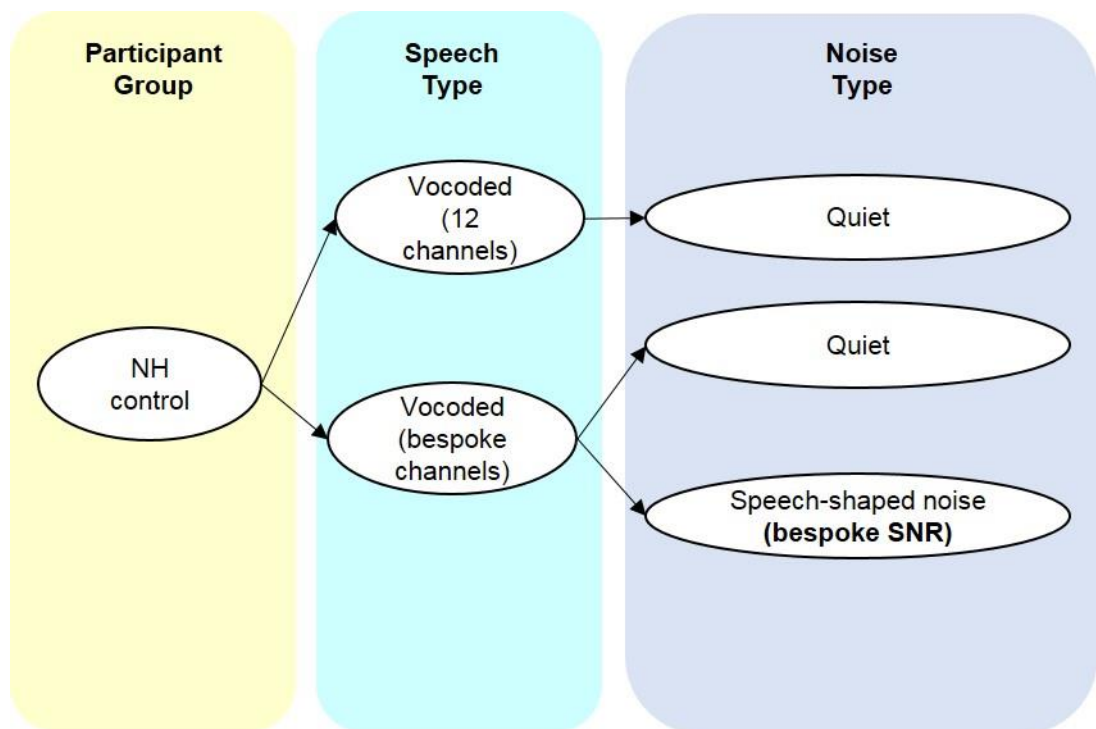


Figure 5.1: A schematic showing the breakdown of conditions involved in the CI simulation training, divided according to participant group, speech type, and noise type.

Adaptive tracking (for number of channels and noise level): How the tracking was implemented (and stimuli used) in Experiment 4 was replicated for the noise tracking with only the following modifications:

- The targeted level of auditory accuracy was 37.5% (not 50%)
- The initial step size was 6.5dB (not 10dB) until the first reversal, at which point, the step size became 3dB for the remainder of tracking.
- The adaptive algorithm was now that, if the participant typed fewer than 2 key words correctly (out of 4), the SNR would increase on the next trial (thus becoming easier), and if the participant successfully heard (and typed) 2 or more key words, the SNR would decrease (and become more difficult).
- Length of track was extended to 25 sentences.

The same software used for the noise tracking was also utilised to implement the channel tracking. This time, however, channel number was manipulated instead of the SNR. The objective of the channel tracking was to find the spectral resolution that elicited 75% accuracy in performance in quiet. The starting channel number was at a very high spectral resolution of 24 channels.

The number of channels then either divided, or multiplied, by 2 until the first reversal. At this point, it then divided, or multiplied, by 1.67 until the second reversal. From the second reversal onwards, the step size was a division or multiplication of 1.4.

After each division or multiplication, the channel number was rounded up or down to the nearest whole number. A maximum of 20 BEL sentences were used within this channel tracking.

The selection of channel number during the tracking (i.e. the adaptive algorithm) was such that, if the participant had typed fewer than 3 key words correct, the channel number would increase (i.e. the channel number would multiply by the requisite step size, and then be rounded up or down to the nearest whole number).

If the participant successfully heard (and typed) 3 key words, the channel number remained the same. If the participant heard all 4 key words correctly, the channel number would decrease (i.e. the channel number would divide by the requisite step size, and then be rounded). The participant's bespoke number of channels was derived by finding the mean of all levels visited (in channel number) at the final even number of reversals.

75% auditory accuracy was chosen for the channel tracking to ensure the participants were not at ceiling performance in quiet (as this would be unrepresentative of typical CI performance).

37.5% accuracy was chosen for the noise tracking in an attempt to ensure a sufficiently large change in acoustic challenge for all participants when noise was introduced. Although this may seem to be a low level of performance, in terms of ecological validity, there were two reasons for its implementation.

Firstly, performance level would already have been adjusted to 75% in quiet, so to maximise the likelihood of detecting an effect when noise was introduced, acoustic challenge needed to be high. Secondly, the low performance level was used in an attempt to ensure the NH controls' behaviour resembled more closely the CI user population. This is particularly important because in this experiment, participants were processing the CI sound through a hearing brain, rather than the deafened brain of a typical CI user. The challenge for the NH controls, therefore, had to be increased in an attempt to mimic this extra disadvantage of the CI user.

As before in Experiment 4, psychometric functions were obtained for both types of tracking.

Test conditions of the Digit Stream Dual-Task (and trial structure): In total, there were four test conditions attempted (Figure 5.2), of which there was one speech type (i.e. vocoded auditory stimuli), two conditions for task type (i.e. auditory-only or dual-task), as well as two conditions for noise type (i.e. quiet or noise). There were 12 trials for each test condition. The trial structure remained the same as in Experiment 4.

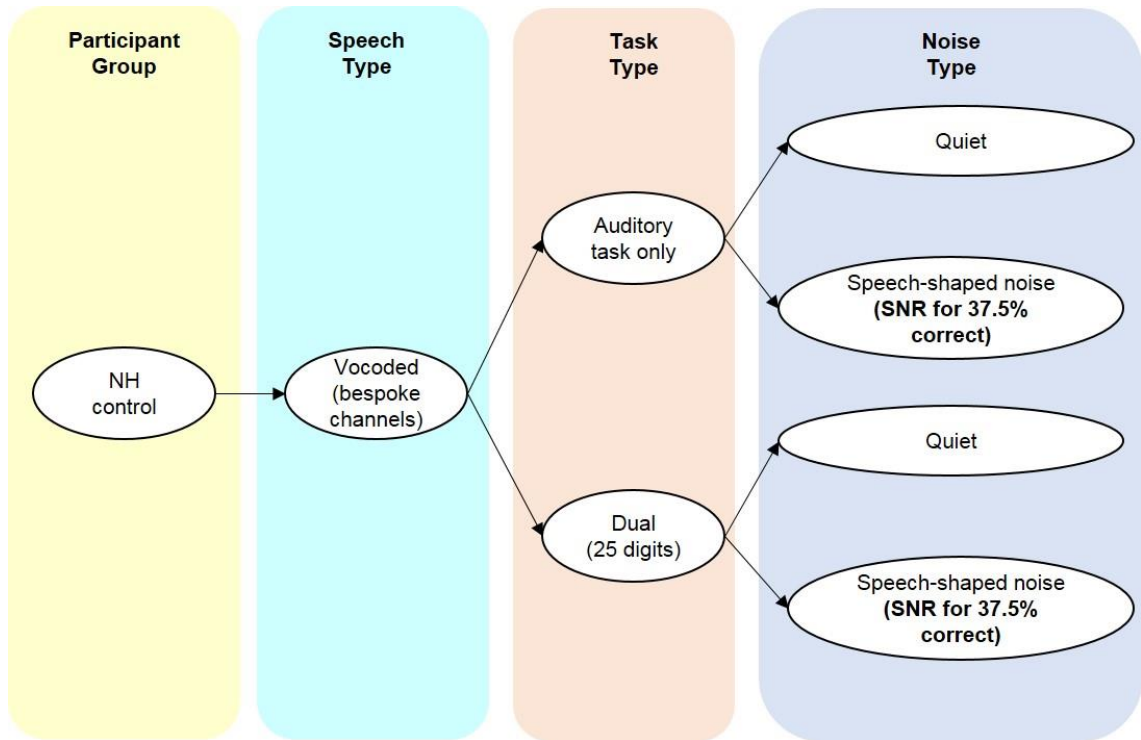


Figure 5.2: A schematic of test conditions (same structure as Figure 3.1).

Protocol: One testing session lasting 90 minutes was conducted, of which 15 minutes was training. For overall protocol, see Table 5.1. All test conditions of the Digit Stream Dual-Task were counterbalanced and randomised across participants in an attempt to counteract training and/or learning effects (i.e. steps 8-11 in Table 5.1).

Before the test run began, the NH controls also had a practice run where they were provided 5 trials of the visual task on its own (presented in quiet) and 5 trials for the dual-task condition.

Statistical analysis: Due to the small sample size, only descriptive statistics were pursued.

5.3. Experiment 5

Table 5.1: Protocol for Experiment 5. Same detail provided as for Table 2.3, including red numbers (with asterisks) indicating all the steps in the protocol where the order of the test conditions were randomised and counter-balanced together.

Step	Training, Practice or Test?	Task Type or Test Condition	Speech Type	Noise Type	Total number of trials/ Total duration
1	Training	Indiana Jones (Crystal Skull)	Vocoded (12 channels)	Quiet	5 minutes
2	Test	Adaptive tracking (tracking 75% correct)	Vocoded (variable channels)	Quiet	20 trials
3	Training	Indiana Jones (Crystal Skull)	Vocoded (bespoke channels)	Quiet	5 minutes
4	Training	Indiana Jones (Crystal Skull)	Vocoded (bespoke channels)	Speech-shaped noise at 10dB SNR	5 minutes
5	Test	Adaptive tracking (tracking 37.5% correct)	Vocoded (bespoke channels)	Speech-shaped noise at variable SNR	25 trials
6	Practice	Visual task only at 25 digits	N/A	Quiet	5 trials
7	Practice	Dual at 25 digits	Vocoded (bespoke channels)	Speech-shaped noise at SNR for 37.5% correct	5 trials
8 *	Test	Auditory task only	Vocoded (bespoke channels)	Quiet	12 trials
9 *	Test	Auditory task only	Vocoded (bespoke channels)	Speech-shaped noise at SNR for 37.5% correct	12 trials
10 *	Test	Dual at 25 digits	Vocoded (bespoke channels)	Quiet	12 trials
11 *	Test	Dual at 25 digits	Vocoded (bespoke channels)	Speech-shaped noise at SNR for 37.5% correct	12 trials

5.3.3. Results and Discussion: Descriptive statistics

5.3.3.1. *LE measurement (secondary task performance): Visual accuracy*

When the visual performance of the secondary task between dual-tasking in quiet and dual-tasking in noise was compared, the majority of the group appeared to improve in accuracy when there was noise in the background (Figure 5.3 a). Therefore, LE appeared to either be unchanged, or decrease, in noise (Figure 5.3 b). These findings were opposite to the original hypothesis.

There are three possible explanations for why negative LE scores were obtained. The first possibility was that the negative LE scores reflected a genuine decrease in LE expended. However, this is improbable since it was not just noise being presented, there was also vocoding (i.e. spectral resolution had been compromised). Therefore, the acoustic challenge presented by the stimulus was likely to be high, meaning an increase in LE would be expected (not a decrease).

The second possibility is that resource allocation executed by the participant had become awry during dual-tasking, with the secondary task being prioritised instead of the primary task (as instructed). Therefore, more cognitive resources could have been allocated towards visual performance, hence the bolstering of visual accuracy observed. Since the participants were very clearly instructed to prioritise the primary task at all times, this incorrect prioritisation (if it had occurred) may very well have been unconscious (or even automatic) and, thus, inadvertent.

The third potential explanation is that the digit stream dual-task was unable to measure LE in the first instance, so these LE scores were essentially meaningless.

To ascertain which of these three explanations were the most likely, the subjective ratings were explored in more detail.

5.3. Experiment 5

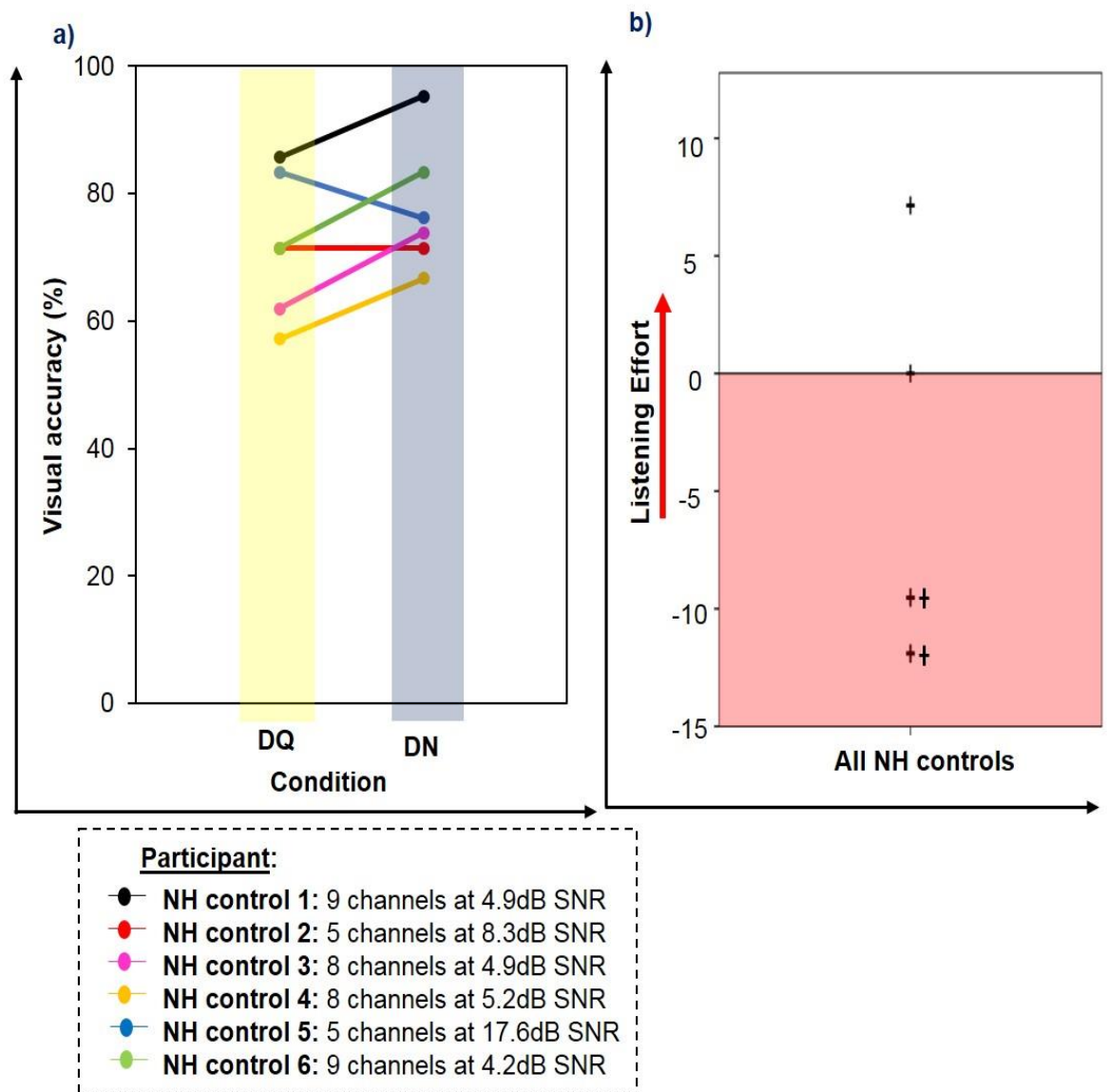


Figure 5.3: Dot plots of visual accuracy (a) and LE (b).

DQ = Dual-tasking in quiet, DN = Dual-tasking in noise. Red arrow indicates direction in which LE increases. Red shading indicates when LE has decreased in presence of noise.

(N.B. For all test conditions conducted in quiet, the auditory stimuli were vocoded with the participants' bespoke number of channels. For all test conditions conducted in noise, the auditory stimuli were vocoded with the participant's bespoke number of channels and the noise level was set at the participants' bespoke SNR.)

Unit of LE = Difference in percentage.

5.3. Experiment 5

5.3.3.2. Evaluation of validity of LE measurement: Subjective ratings.

There was the general trend of perceived effort becoming higher (i.e. towards that of “Extremely hard work”) when NH controls transitioned from quiet to noise, especially when dual-tasking (Figure 5.4). Unlike the secondary task performance, these subjective ratings were precisely as hypothesised. Thus, the experience of perceived effort increased when participants performed the dual-tasking in noise, which was inconsistent with what was indicated by visual performance.

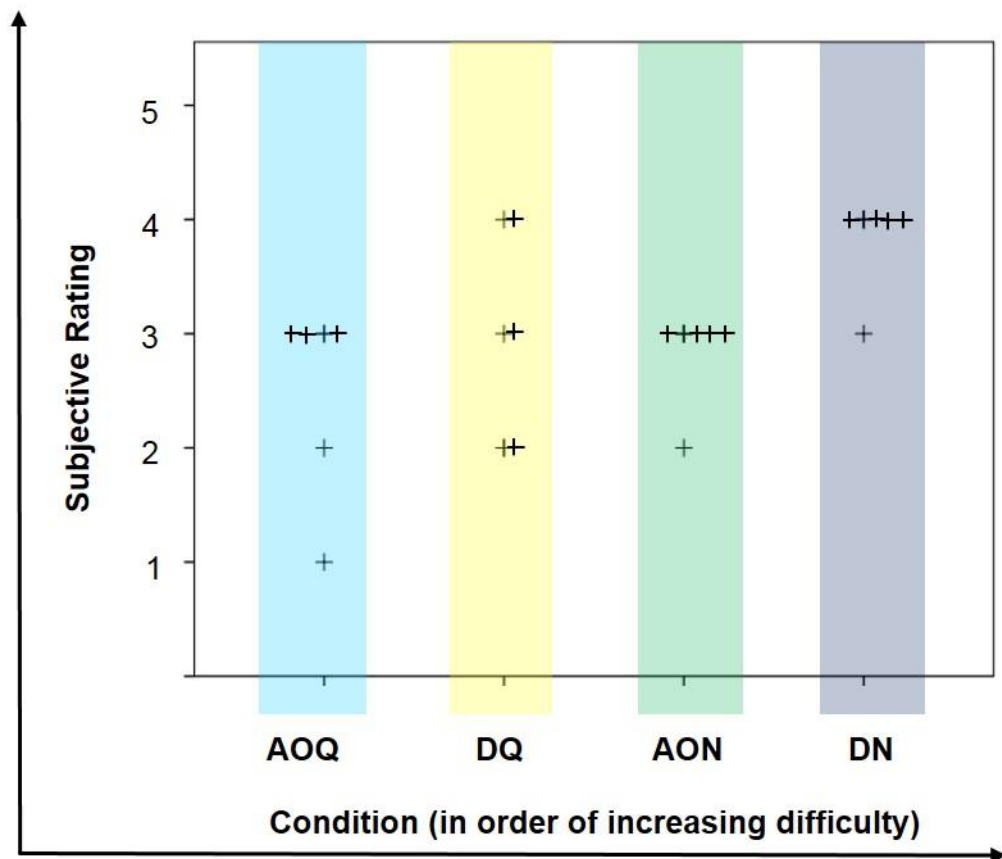


Figure 5.4: Dot plot of subjective ratings.

AOQ = Auditory task only in quiet, DQ = Dual-tasking in quiet, AON = Auditory task only in noise, DN = Dual-tasking in noise.

Subjective ratings: 1 = Not at all hard work, 2 = Quite hard work, 3 = Medium hard work, 4 = Very hard work, 5 = Extremely hard work.

The lack of subjective corroboration gives rise to the possibility that there may have been inappropriate resource allocation within the execution of the digit stream dual-task. To verify this, primary task stability needs to be calculated.

5.3.3.3. *Stability of primary task performance:*
Auditory accuracy

It was found that there was an overall tendency for auditory accuracy to deteriorate when noise was introduced in both single-task and dual-task conditions (Figure 5.5 a; see Figure 5.5 b for applied testing parameters for auditory stimuli). When primary task stability was specifically calculated (Figure 5.5 c), it emerged that deterioration had also occurred when dual-tasking was required in noise.

Because the participants had been tested in quiet (as well as noise), for both the auditory-only and dual-tasking conditions, it was now possible to ascertain primary task stability for the quiet listening condition too. This was particularly important because there was now acoustic challenge present within the quiet listening condition (i.e. the vocoding), and its impact on auditory accuracy needs to be understood. It emerged that a similar number of participants deteriorated in auditory performance when dual-tasking in quiet too, although the extent of this deterioration was not as marked (compared to that in noise).

These two types of deterioration have the following implications:

- The acoustic challenge in the form of noise level (combined with vocoding) may have encouraged the tendency to favour the secondary task when dual-tasking, hence the negative LE scores;
- The vocoding alone may have produced sufficient acoustic challenge to cause inappropriate resource allocation.

5.3. Experiment 5

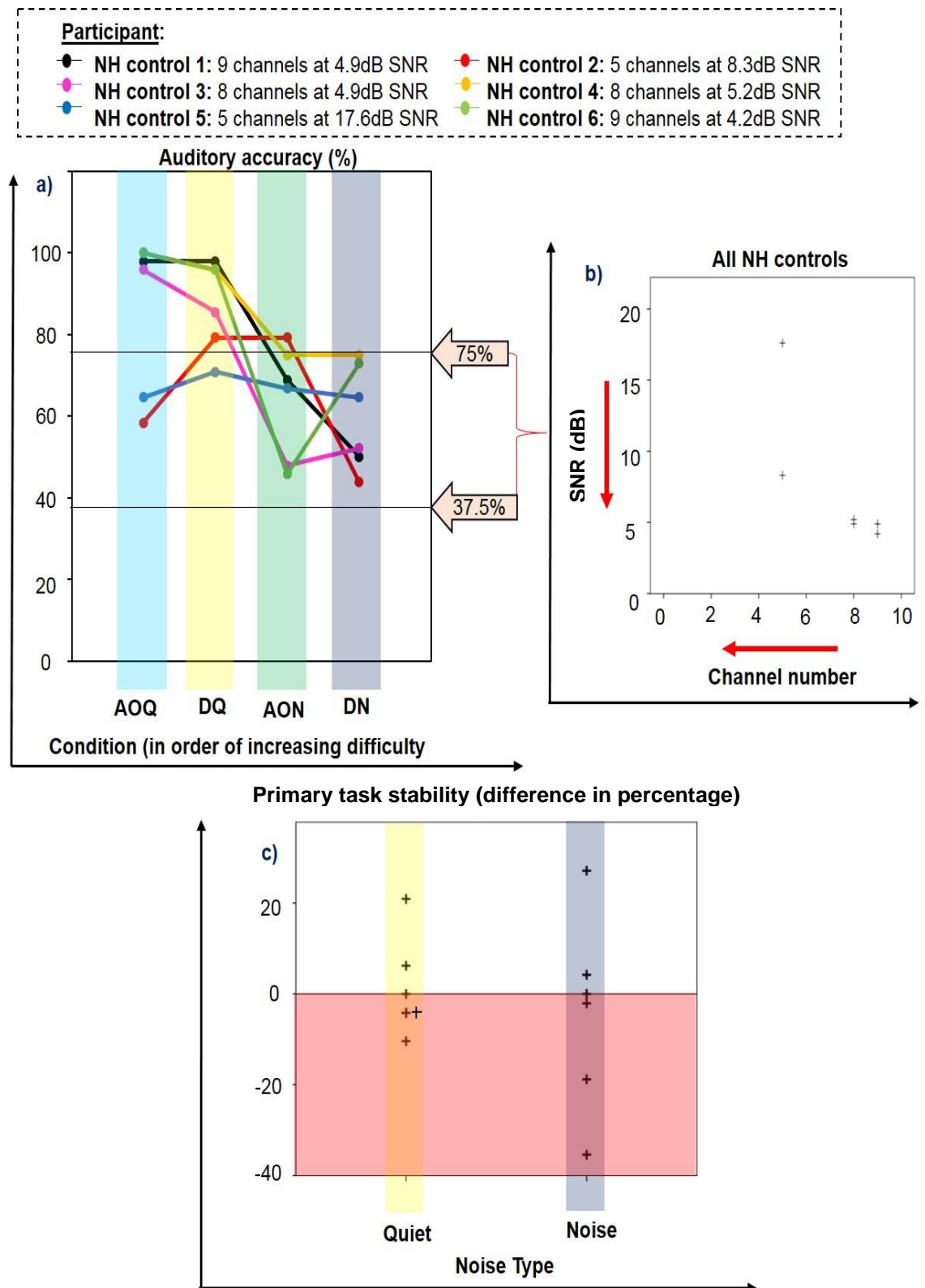


Figure 5.5: Dot plots of auditory accuracy (a) and primary task stability (c), as well as a scatterplot of applied testing parameters (b)

AOQ = Auditory task only in quiet, DQ = Dual-tasking in quiet, AON = Auditory task only in noise, DN = Dual-tasking in noise. Red arrow indicates direction of superiority in value (b). Red shading indicates deterioration in auditory accuracy when dual-tasking (c).

5.3. Experiment 5

5.3.3.4. Evaluation of viability of LE measurement

To understand the cause of the potential instability in primary task performance, the tracking data was revisited. It was confirmed that reasonable tracks were produced (as well as the expected sigmoidal shapes within the psychometric functions) for the majority of cases within both the channel and noise tracking (for exemplars, see Figure 5.6 c-f; for remainder of NH controls' psychometric functions and matching tracks, see Appendices 5.1-5.4). However, when the participants' actual auditory accuracy during the auditory-only conditions were checked against the intended performance level (i.e. 75% in quiet [AOQ], 37.5% in noise [AON]), all but two NH controls performed better than designed in both instances (Figure 5.6 a & b).

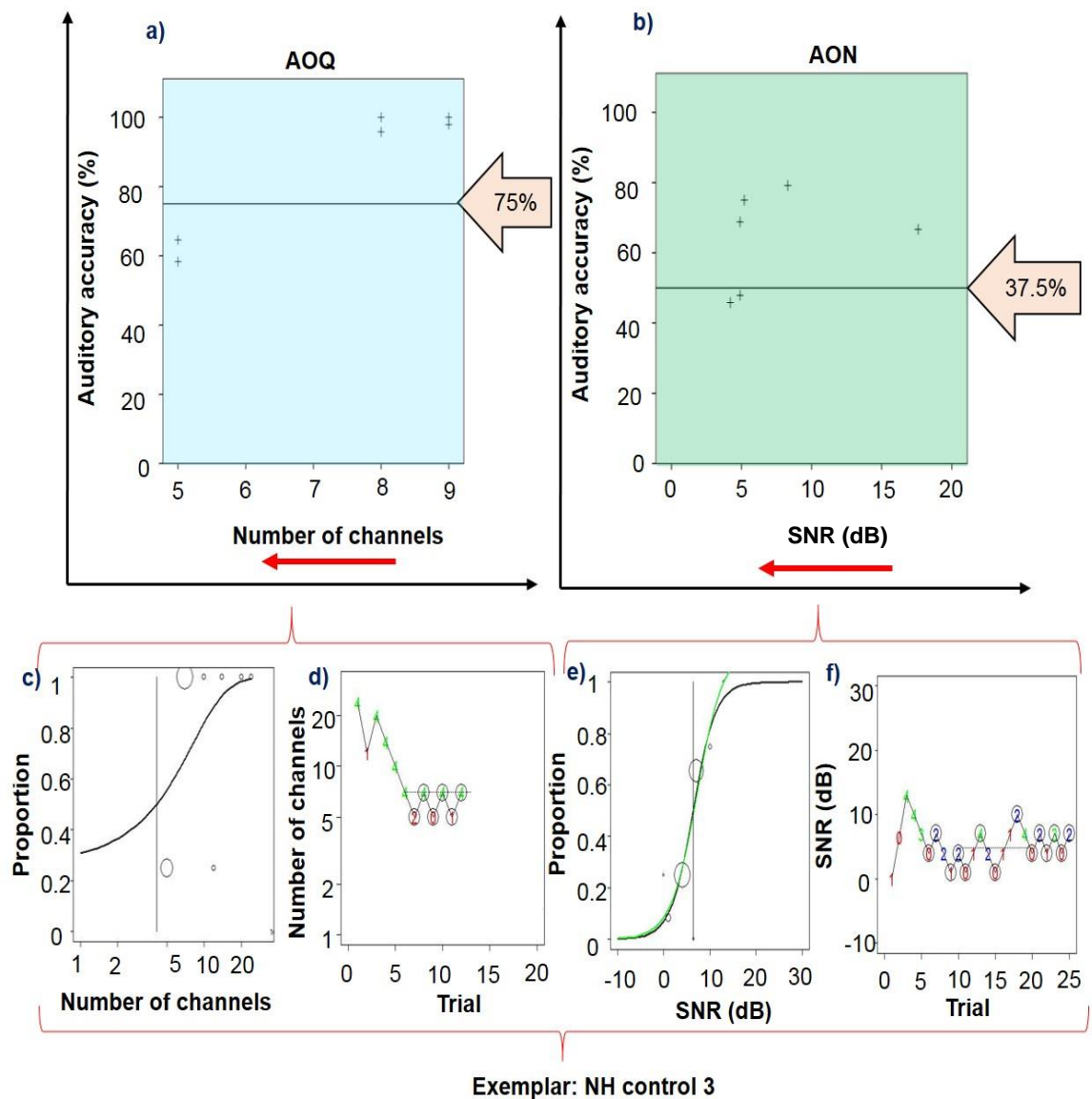


Figure 5.6: Scatterplots of auditory accuracy versus channel number and SNR (a, b). Exemplars of psychometric function and adaptive track for channel and noise tracking (c, d, e, f). Black line on psychometric function = standard logistic regression. Green line on psychometric function = logistic regression with additional parameter in the fit, i.e. an upper asymptote, otherwise known as lapse rate (to enable psychometric function to plateau at $y < 1$).

A potential explanation for this improved auditory accuracy could be the continued perceptual learning of the vocoded stimuli by the participants within the testing of the digit stream dual-task itself. Prior hearing research has cited this perceptual learning to be rapid; so rapid that speech intelligibility scores can improve from floor performance (i.e. 0%) to that of 70% accuracy following exposure to just 30 sentences (Davis et al., 2005). Not only is perceptual learning fast, it is also sufficiently potent to be able to generalise to untrained words too (implying that perceptual learning is not merely rote learning of distorted words, or enhanced guessing: Davis et al., 2005; Hervais-Adelman et al., 2008).

Accordingly, this experiment (i.e. Experiment 5) operated under the premise that perceptual learning of vocoded stimuli would have been achieved quickly enough that testing of the digit stream dual-task with CI simulations could also be conducted within the same session as the training. Although this was not unreasonable to assume, it was still possible that there could have been incomplete transfer of this learning. Indeed, this ongoing transfer of learned stimuli has been previously demonstrated in CI simulations research, especially when more complex acoustic features are involved, such as those present in speech (Hervais-Adelman et al., 2011). This ongoing transfer of perceptual learning would have the consequence of bolstering auditory accuracy within the digit stream dual-task (i.e. after training was completed by the participant).

Ultimately, the improvement in auditory performance gives rise to the possibility that the derived values for spectral resolution (i.e. the channel number) and SNR were not actually representative of the participants' ability. Therefore, it seems unlikely that the two types of tracking achieved their objective of equalising auditory performance level in both quiet and noise. This, in turn, may account for the primary task instability.

To further exacerbate matters, there was also the recurrence of over-reporting. When these instances of over-reporting were isolated (Figure 5.7a), it was discovered that 4 out of the 6 NH controls were responsible (with over-reporting occurring in a maximum of 2 trials). However, this over-reporting only emerged when there were 1 or 2 targets presented in the digit stream. Otherwise, the nature of visual response was as expected: as more targets were presented in the digit stream, the participant tended to report seeing fewer targets (as revealed when the different type of visual responses were tallied up for each target presentation frequency in each test condition: Figure 5.7 b-g).

The combination of primary task instability and over-reporting gives rise to the possibility that LE measurement is no longer reliable.

5.3. Experiment 5

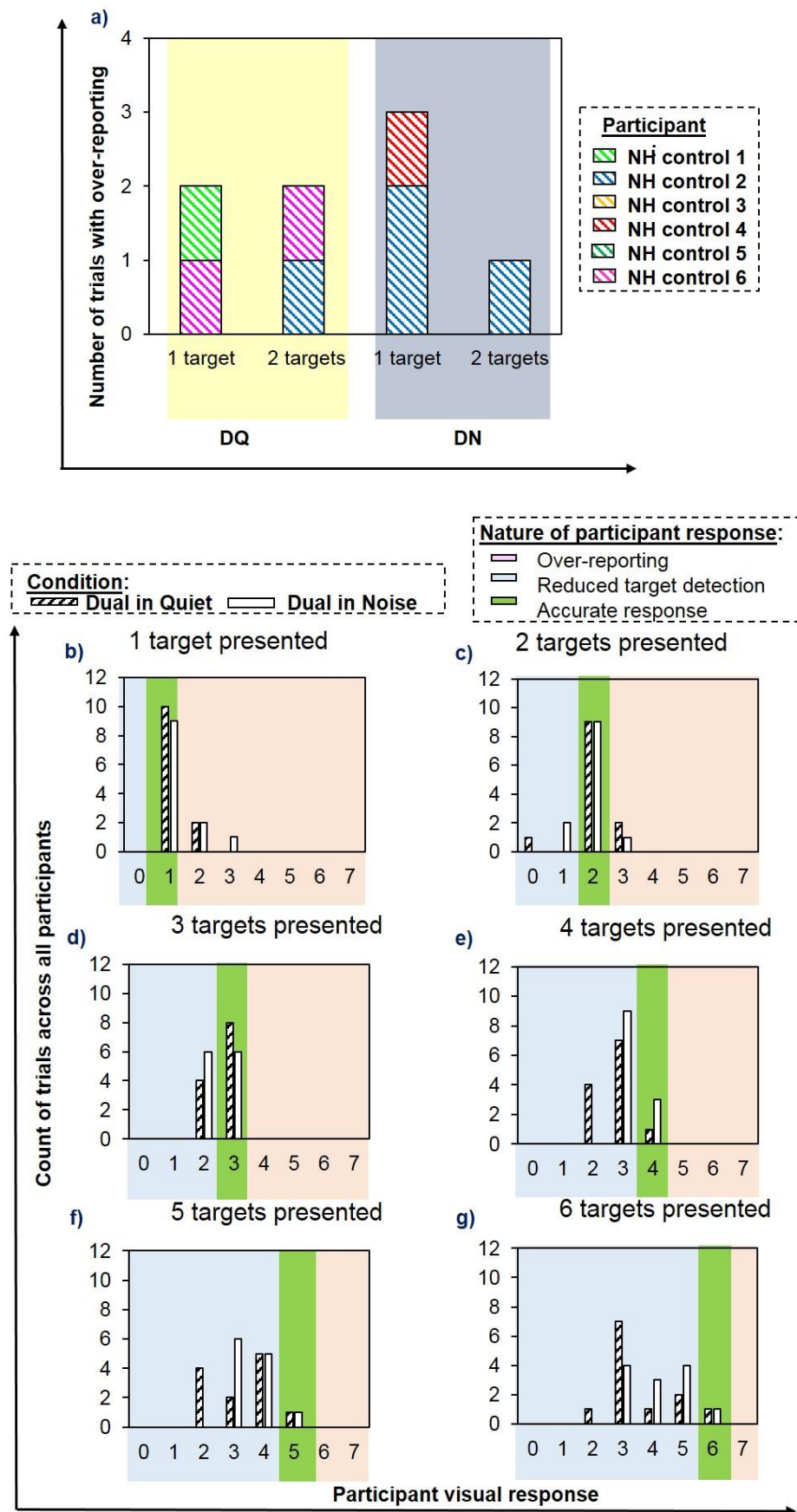


Figure 5.7: Stacked histogram showing the incidence of over-reporting (a), and histograms depicting the nature of visual responses for all participants across all trials (in each test condition) for each target presentation frequency (b, c, d, e, f, g).

5.4. Conclusions of Experiment 5

In this first attempt of incorporating CI simulations within the digit stream dual-task (in order to test this LE measurement in listening conditions more closely resembling the electric hearing produced by the CI neuroprosthesis), counter-intuitive LE scores were obtained. Not only this, there was instability within the primary task performance as well as persistence of over-reporting within secondary task performance. It is possible that the underlying cause of this aberrant behaviour was ultimately the continued transfer of perceptual learning of the vocoded stimuli (after training was completed). Therefore, the acoustic challenge presented was not as intended which may then have skewed resource allocation, thus compromising LE measurement.

5.5. Proposed solution

Before definitively concluding that the digit stream dual-task is an unviable methodology for LE measurement, there should be an attempt to ameliorate any potential interference of perceptual learning. After all, if it was indeed this perceptual learning compromising the LE measurement, then no blame lies within the design of the digit stream dual-task itself. Accordingly, another small-scale study with a prolonged training session is pursued, in order to confirm the testing parameters required for a larger-scale study.

To enable ample time for training, two data collection sessions are constructed for Experiment 6: one devoted exclusively to CI simulation training, and the other for LE testing itself. Because the training session is now 90 minutes long, there is now the opportunity to be more thorough in the level of spectral resolution the participant is exposed to, i.e. there is now more gradation in channel number applied in the noise-excited vocoding within the steps of training.

Not only will training time be extended, but a minimum of 24 hours (which includes at least one night of sleep), will also be required as an interval between the training and testing sessions. This is in an attempt to control for any potential sleep consolidation effects on learning and memory. This is because rapid eye movement (i.e. REM) has been postulated to induce an increase of neuronal plasticity following an enriched (or novel) waking experience, thus establishing and consolidating newly acquired information within the hippocampal and cortical (e.g. prefrontal and parietal) regions in the brain (Ribeiro et al., 1999; Walker, 2008). Although there is controversy within this research field, sleep effects still warrant factoring into the experiment to try to minimise any other types of disruption to the perceptual learning and its transfer (Vertes, 2004).

In addition, the criterion used for the noise tracking will be eased from that used in Experiment 5. This is to compensate for any possibility that the added acoustic of noise combined with the vocoding would overwhelm the participant to their detriment (especially within the dual-tasking conditions). This is in light of the observation of primary task instability being possible even in quiet, implying that coping with the vocoded stimuli alone is already difficult.

5.6. Experiment 6

5.6.1. Hypotheses

The extension to training time will ameliorate any interference from perceptual learning. Therefore, the use of the two adaptive tracking will succeed in equalising the auditory performance levels in both quiet and noise. Accordingly, LE will now increase when noise is introduced, i.e.

- Visual accuracy will decrease when the SNR decreases.

This change in visual accuracy will not be at the cost of resource allocation:

- Auditory accuracy will remain constant when the participant is dual-tasking (relative to that achieved when performing only the auditory task).
- Over-reporting will also be ameliorated.

5.6.2. Methods

5.6.2.1. Participants

10 adult normal hearing (NH) controls (3 females and 7 males, mean age: 54.7 years, ± 13.82 S.D) were recruited. All participants had normal, or corrected-to-normal vision, and were native British English speakers.

5.6.2.2. Procedure

Location of testing: Testing sessions took place within the participants' homes.

Primary task of the Digit Stream Dual-Task: Same as before.

Secondary task of the Digit Stream Dual-Task: Same as before.

Acoustic challenge: Same as before.

Vocoding parameters: As for Experiment 5.

CI simulation training: This training amounted to 90 minutes in total. The same training software and stimuli used in Experiment 5 was implemented. There were 6 conditions of spectral resolution (with and without speech-shaped noise in the background) that the participant was exposed to (Figure 5.8).

This involved 4 conditions in channel number applied within the vocoding (i.e. the number of analysis filters, or frequency bands, implemented in the auditory stimuli): 12 channels, 8 channels, 6 channels, and also the participant's bespoke channel number. This bespoke channel number was obtained via adaptive tracking (of which there were two attempts: for more detail, see section about adaptive tracking).

When the spectral resolution was set at the participant's bespoke channel number, speech-shaped noise was then introduced, with this training involving three conditions in SNR: 10dB SNR, 7dB SNR, and also the participant's SRT. The participant's SRT was obtained via adaptive tracking (of which there were two attempts: for more detail, refer to section about adaptive tracking).

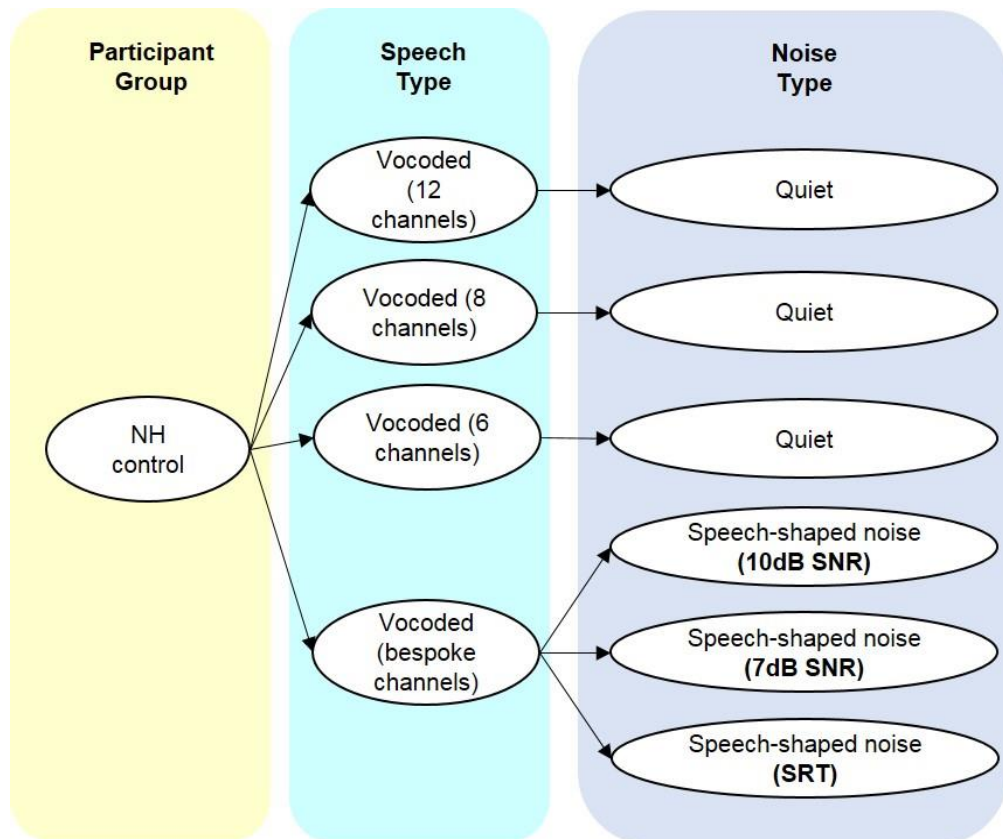


Figure 5.8: A schematic of conditions within CI simulation training (same structure for Figure 5.1).

5.6. Experiment 6

For overall protocol of training (which includes when the two types of adaptive tracking were conducted), see Table 5.2. The difficulty level was systematically increased as the participant progressed through the training, firstly with the reduction in spectral resolution, and then with the introduction of noise (with SNR steadily worsening). The training session was then concluded with the opportunity to practice the Digit Stream Dual-Task before the actual testing session was begun. This practice involved 10 trials of just the secondary visual task on its own in quiet, as well as 18 trials of the dual-task condition of the Digit Stream Dual-Task, with the auditory stimuli vocoded at 12 channels and no background noise present.

Table 5.2: Protocol for training in Experiment 6. Same detail provided regarding box shading as for Table 2.3. Please note no counter-balancing or randomisation apply in this protocol (i.e. all steps are executed in fixed order as listed).

Step	Training, Practice or Test?	Task Type or Test Condition	Speech Type	Noise Type	Total number of trials/ Total duration
1	Training	Indiana Jones	Vocoded (12 channels)	Quiet	2 minutes
2	Training	Indiana Jones	Vocoded (8 channels)	Quiet	5 minutes
3	Training	Indiana Jones	Vocoded (6 channels)	Quiet	5 minutes
4	Test	Adaptive tracking (tracking 75% correct)	Vocoded (variable channels)	Quiet	20 trials
5	Test	Adaptive tracking (tracking 75% correct)	Vocoded (variable channels)	Quiet	20 trials
6	Training	Indiana Jones	Vocoded (bespoke channels)	Speech-shaped noise at 10dB SNR	5 minutes
7	Training	Indiana Jones	Vocoded (bespoke channels)	Speech-shaped noise at 7dB SNR	5 minutes
8	Test	Adaptive tracking (tracking 50% correct)	Vocoded (bespoke channels)	Speech-shaped noise at variable SNR	20 trials
9	Training	Indiana Jones	Vocoded (bespoke channels)	Speech-shaped noise at SRT	5 minutes
10	Test	Adaptive tracking (tracking 50% correct)	Vocoded (bespoke channels)	Speech-shaped noise at variable SNR	20 trials
11	Practice	Visual task only at 25 digits	N/A	Quiet	10 trials
12	Practice	Dual at 25 digits	Vocoded (12 channels)	Quiet	18 trials

Adaptive tracking (channel and noise): The overall protocol (and procedures) involved for both channel and noise tracking were the same as those used in Experiment 5, with only the following changes:

- The targeted level of auditory accuracy for the noise tracking was now 50% correct, i.e. the participant's SRT (not 37.5%).
- The starting SNR for the noise tracking was 20dB.
- Both types of tracking were executed twice.

Consequently, the first step size was 10dB until the first reversal and 6.5dB until the second reversal, at which point, step size was fixed 3dB from that point.

In addition, the adaptive algorithm was the following: if a participant achieved comprehension of less than 2 key words, the SNR increased on the next trial; if the participant typed 2 key words correctly, the SNR remained the same; and if more than 2 key words were typed, the SNR decreased on the next trial.

If a lower spectral resolution or SRT was achieved in the second channel tracking, this was utilised as the participant's bespoke channel number for the remainder of the training (and also for testing in the next session). If better performance was obtained on the first run, a rounded average was taken of the values produced by the two tracks in order to derive the participant's bespoke testing parameters.

Test conditions of Digit Stream Dual-Task (and trial structure): In total, there were five test conditions attempted (Figure 5.9), of which there was one speech type (i.e. vocoded auditory stimuli, set at the participant's bespoke number of channels), three conditions for task type (i.e. visual-only, auditory-only, or dual-task), and two conditions for noise type (i.e. quiet or noise with SNR set at the participant's SRT).

There were 18 trials for each test condition. The trial structure remained the same as before.

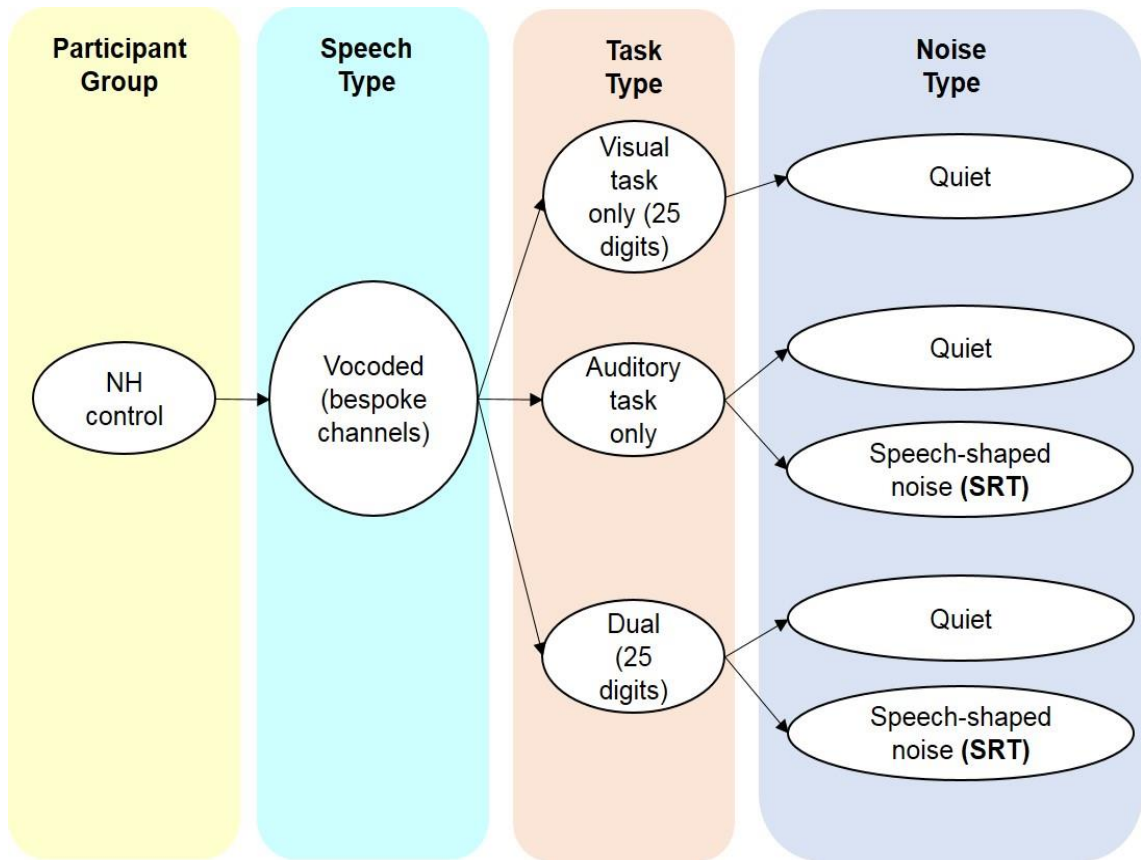


Figure 5.9: A schematic of test conditions (same structure as Figure 3.1).

Protocol: The session for the testing of the Digit Stream Dual-Task was conducted after the mandatory minimum of 24 hours rest, which included at least one night of sleep.

For overall protocol of the testing session, see Table 5.3. All test conditions of the Digit Stream Dual-Task were counterbalanced and randomised across participants in an attempt to counteract training and/or learning effects (i.e. steps 6-10 in Table 5.3). Before the test run began, the NH controls had the opportunity to warm up via 5 minutes of the CI simulation training protocol conducted in noise, where the spectral resolution was set at the participant's bespoke channel number (as obtained from the two attempts at channel tracking during the training session), and noise being set at the participant's SRT (as obtained from the two attempts at noise tracking during the training session).

5.6. Experiment 6

In addition, the participant had a practice run of the Digit Stream Dual-Task, where they executed 10 trials of the visual task on its own (presented in quiet), 18 trials for the dual-task condition (with spectral resolution set at 12 channels and no background noise), 18 trials for the dual-task condition again (with same spectral resolution of 12 channels but with noise with SNR at the participant's bespoke SRT), and also 18 trials for the dual-tasking condition yet again (with spectral resolution now at the participant's bespoke channel number as well as noise being present with the same SNR, i.e. at the participant's SRT).

Statistical analysis: Due to the small sample size, only descriptive statistics were pursued.

Table 5.3: Protocol for testing of the Digit Stream Dual-Task in Experiment 6. Same detail provided as for Table 2.3, including red numbers (with asterisks) indicating all the steps in the protocol where the order of the test conditions were randomised and counter-balanced together.

Step	Training, Practice or Test?	Task Type	Speech Type	Noise Type	Total number of trials/ Total duration
1	Training	Indiana Jones	Vocoded (bespoke channels)	Speech-shaped noise at SRT	5 minutes
2	Practice	Visual task only at 25 digits	N/A	Quiet	10 trials
3	Practice	Dual at 25 digits	Vocoded (12 channels)	Quiet	18 trials
4	Practice	Dual at 25 digits	Vocoded (bespoke channels)	Quiet	18 trials
5	Practice	Dual at 25 digits	Vocoded (bespoke channels)	Speech-shaped noise at SRT	18 trials
6 *	Test	Auditory task only	Vocoded (bespoke channels)	Quiet	18 trials
7 *	Test	Auditory task only	Vocoded (bespoke channels)	Speech-shaped noise at SRT	18 trials
8 *	Test	Dual at 25 digits	Vocoded (bespoke channels)	Quiet	18 trials
9 *	Test	Dual at 25 digits	Vocoded (bespoke channels)	Speech-shaped noise at SRT	18 trials
10 *	Test	Visual task only at 25 digits	N/A	Quiet	18 trials

5.6.3. Results and Discussion: Descriptive statistics

5.6.3.1. LE measurement (secondary task performance): Visual accuracy

Overall, there was variability in level of visual accuracy in the NH controls, in both quiet and noise, and also across practice and test runs (Figure 5.10 a). The most marked trends were the deterioration of visual accuracy when dual-tasking was required, and also when the spectral resolution of the vocoding reduced to the participant's bespoke channel number (i.e. down from 12 channels). When visual performance when dual-tasking in quiet (with bespoke channel number) was specifically compared to that achieved when dual-tasking in noise, only 4 NH controls elicited positive LE scores when noise was introduced (Figure 5.10 b). The rest either showed no change in LE, or achieved negative scores.

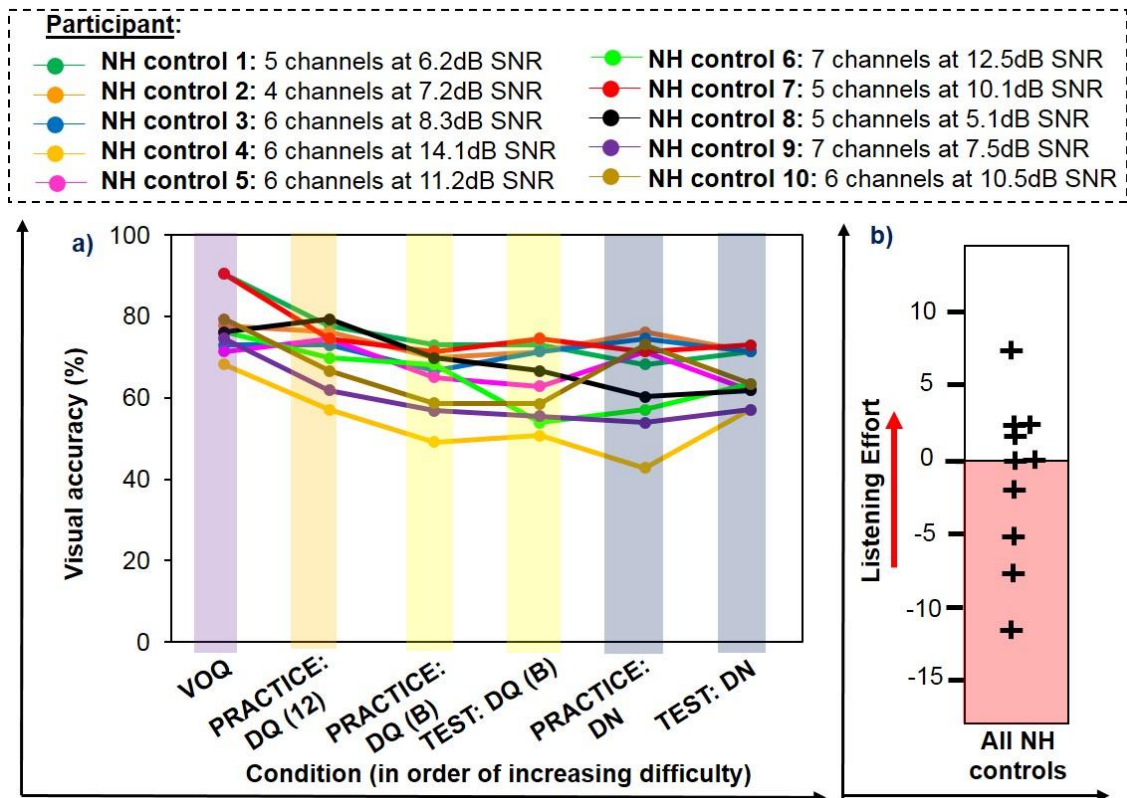


Figure 5.10: Dot plots of visual accuracy (a) and LE (b). Unit of LE = Difference in percentage.

VOQ = Visual task only (in quiet), DQ (12) = Dual-tasking in quiet, with spectral resolution of vocoding set at 12 channels, DQ (B) = Dual-tasking in quiet, with spectral resolution of vocoding set at the participant's bespoke number of channels (as obtained from the two attempts of channel tracking during training session), DN = Dual-tasking in noise with spectral resolution of vocoding set at the participant's bespoke number of channels and SNR set at the participant's SRT (as obtained from the two attempts of noise tracking during the training session). Red arrow indicates direction in which LE increases. Red shading indicates when LE has decreased in presence of noise.

5.6.3.2. *Stability of primary task performance:*
Auditory accuracy

It was found that, in both the practice and test runs, the introduction of noise tended to cause a reduction in auditory accuracy, especially when spectral resolution was at the participant's bespoke channel number (Figure 5.11 a; for testing parameters of auditory stimuli, see Figure 5.11 c). When primary task stability was assessed in quiet (with bespoke channel number) and in noise (Figure 5.11 b), the majority of NH controls deteriorated in performance when required to dual-task in noise. However, not as many NH controls exhibited primary task instability in quiet. As any primary task instability could be interpreted as inappropriate resource allocation, this now brings the positive LE scores into doubt with regards to their validity.

5.6.3.3. *Evaluation of viability of LE measurement*

To better ascertain the reliability of the LE scores, the target detection percentages were revisited to establish the nature of the visual response elicited. It emerged that over-reporting was still occurring.

When the instances of over-reporting were isolated (Figure 5.12), it emerged that all participants were susceptible to over-reporting. However, it was mainly only one trial of each test condition that was contaminated by the over-reporting. In addition, this over-reporting completely disappeared in all test conditions (irrespective of whether it was the practice or test run) where there was more than 3 targets in the digit stream.

5.6. Experiment 6

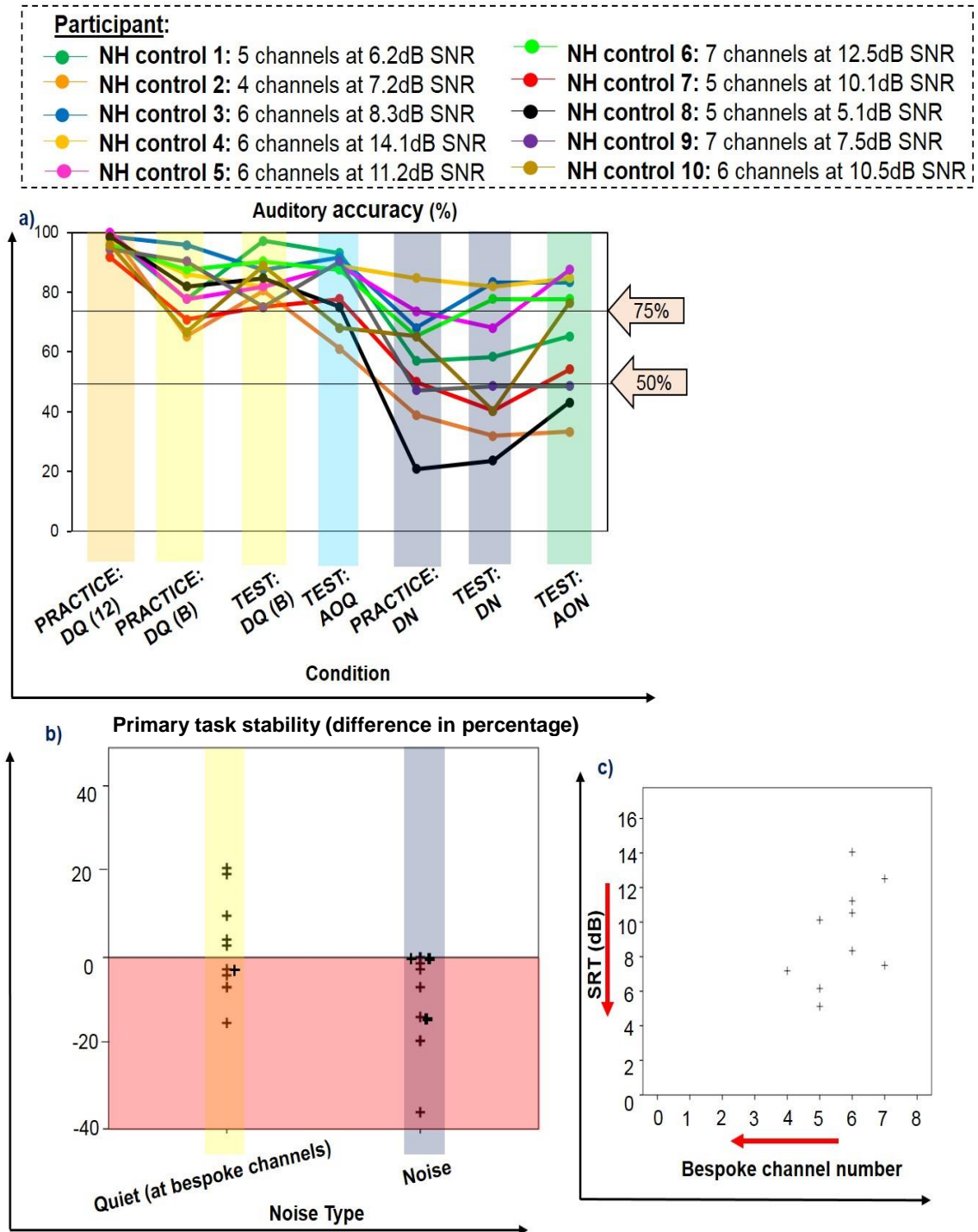


Figure 5.11: Dot plots of auditory accuracy **(a)** and primary task stability **(b)**, and scatterplot of participants' SRT and bespoke channel number **(c)**.

AOQ = Auditory task only in quiet (with spectral resolution of vocoding set at the participant's bespoke channel number), AON = Auditory task only in noise (with spectral resolution of vocoding set at the participant's bespoke channel number and SNR set at the participant's SRT), DQ (B) = Dual-tasking in quiet with spectral resolution set at the participant's bespoke channel number), DQ (12) = Dual-tasking in quiet (with spectral resolution of vocoding set at 12 channels), DN = Dual-tasking in noise (with stimuli vocoded at the participant's bespoke channel number and SNR set at the participant's SRT). Red arrow indicates superiority of value. Red shading indicates deterioration in auditory accuracy when dual-tasking.

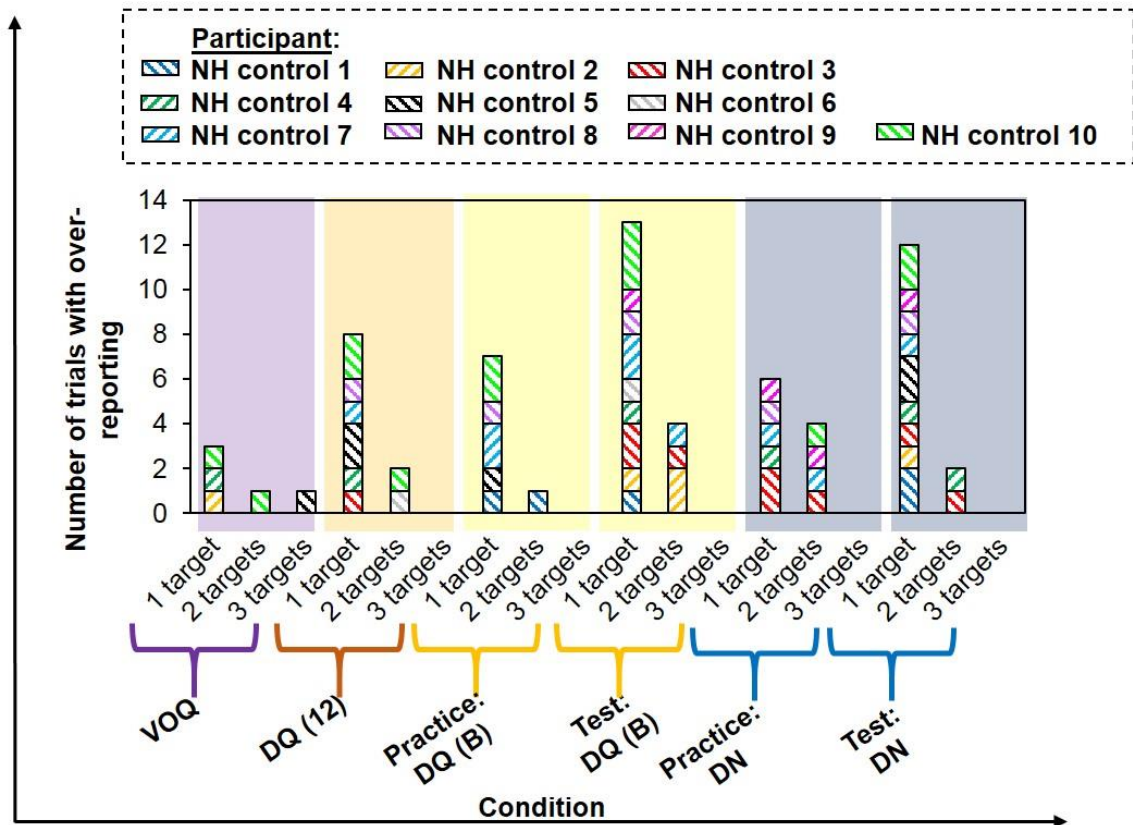


Figure 5.12: Stacked histogram showing the incidence of over-reporting for each participant.

VOQ = Visual task only, DQ (B) = Dual-tasking in quiet with spectral resolution at the participant's bespoke channel number, DQ (12) = Dual-tasking in quiet, with spectral resolution at 12 channels, DN = Dual-tasking in noise with stimuli vocoded at the participant's bespoke channel number and SNR set at the participant's SRT.

Furthermore, the visual response was as expected: more targets were missed as the numbers of targets presented in the digit stream increased (as revealed when the trials of each type of visual response were tallied up across all participants in each test condition for each target presentation condition: Figure 5.13).

However, despite the relative sparseness of over-reporting and visual responses generally being as desired, this does not refute the fact that there was primary task instability. Accordingly, the adaptive tracking data was interrogated to ascertain what might be the underlying cause.

5.6. Experiment 6

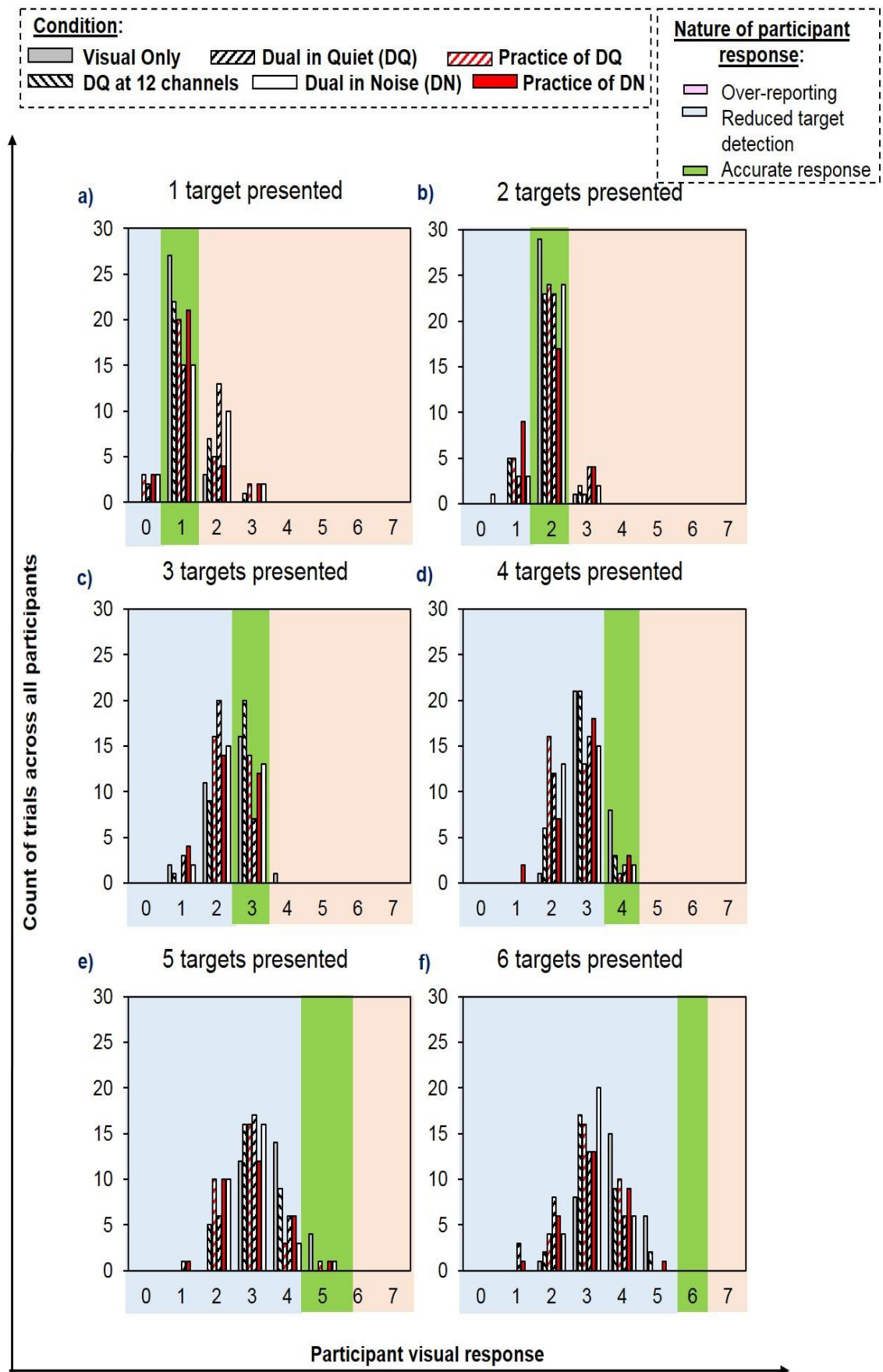


Figure 5.13: Histograms depicting the nature of participants' visual responses across trials (in each test condition) for each target presentation frequency.

The majority of the psychometric functions and affiliated tracks were found to be reasonable across the NH controls, for both the channel and the noise tracking (for exemplars of psychometric function with associated track for the two attempts of both channel and noise tracking, see Figure 5.14 c- k; for remainder of psychometric functions and tracks for NH controls, see Appendices 5.5—5.14). However, when auditory accuracy obtained during the auditory-only conditions in both quiet and noise was plotted as a function of the stimulus parameters used, it was revealed that the majority of NH controls' performances were generally better than the intended 75% accuracy in quiet, and also better than the intended 50% accuracy in noise (for auditory-only performance in quiet plotted as a function of bespoke channel number, see Figure 5.14 a; for auditory-only performance in noise plotted as a function of SRT, see Figure 5.14 b).

This improved auditory performance (relative to that tracked) suggests the possibility that the adaptive tracking did not necessarily succeed in controlling and equalising the participants' performances. In further support of this, there even appeared to be a relationship between SNR and auditory performance, whereby the poorer the SNR, the lower the auditory accuracy became. A similar relationship was found between channel number and auditory performance, but correlation seemed less marked. If the adaptive tracking had truly succeeded in its objective of equalising performance, there would have been no correlation present, with all participants at the same accuracy (in both quiet and in noise).

It needs to be noted that adaptive tracking for noise level was executed after adaptive tracking for channel number (for the vocoding). There is thus the possibility that it is not feasible to obtain a bespoke SNR value (i.e. via tracking) when the participant's performance has already been tracked to be at 75% correct in quiet.

Regardless of the cause of the primary task instability, there is ultimately one conclusion: LE measurement is now likely to be unreliable.

5.6. Experiment 6

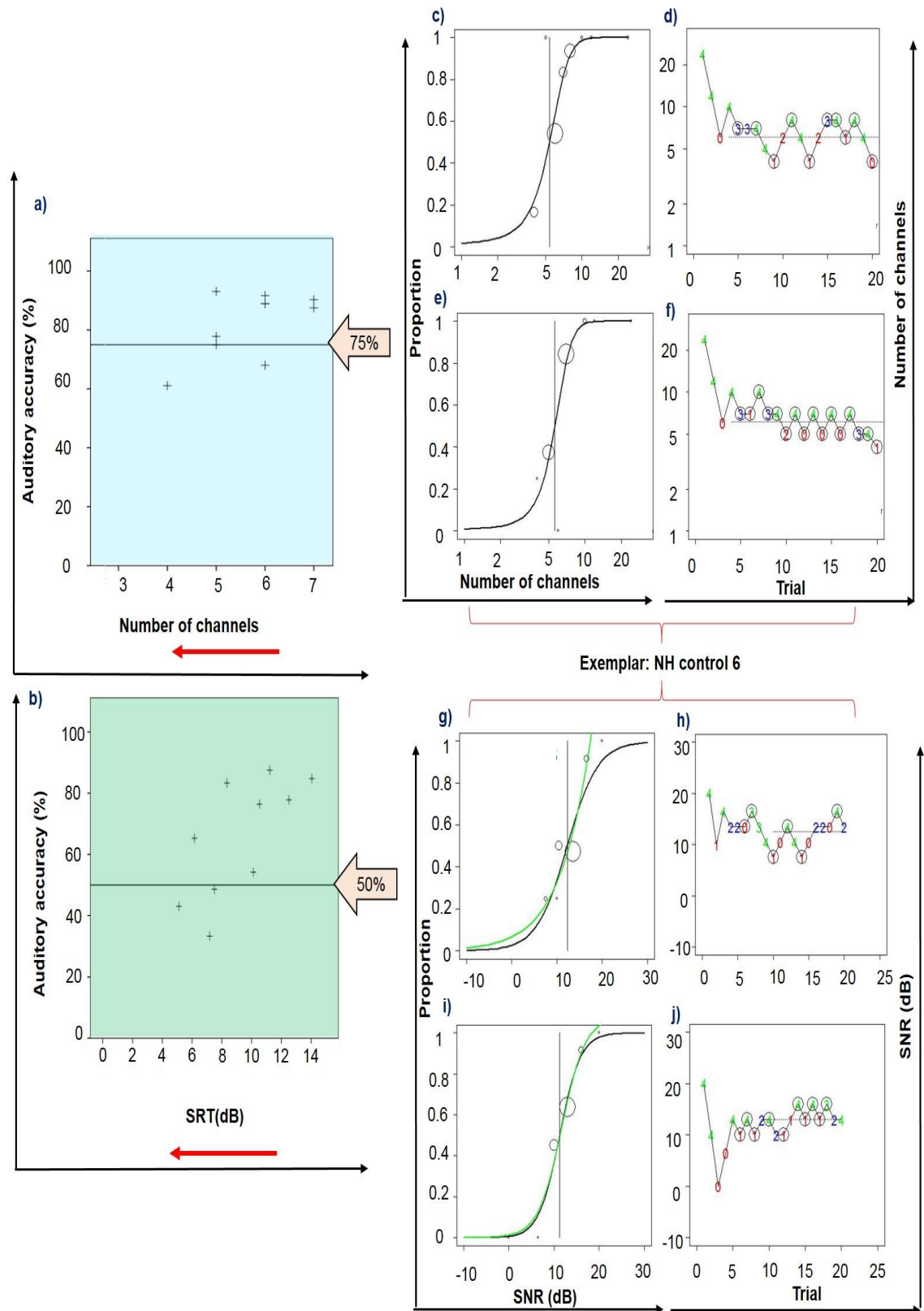


Figure 5.14: Scatterplots of auditory accuracy versus channel number and SNR for all NH controls (a, b) and exemplars of psychometric function and adaptive track (c, d, e, f, g, h, i and j). AOQ = Auditory task only in quiet, AON = Auditory task on its own in noise. Red arrow indicates superiority of score. Black line on psychometric function = standard logistic regression. Green line on psychometric function = logistic regression with additional parameter in the fit, i.e. an upper asymptote, otherwise known as lapse rate (to enable psychometric function to plateau at $y < 1$).

5.7. Conclusions of Experiment 6

In spite of the active attempt to ameliorate the continued perceptual learning effects (regarding the vocoded stimuli) with a separate and extended training session (that was more thorough), there remained primary task instability and also over-reporting (within the secondary visual task). It seems to be that tracking had again failed to equalise auditory performance. In particular, it seems to be the case that trying to execute noise tracking on top of already tracked auditory performance was not a valid form of methodology. Consequently, LE measurement has potentially become compromised.

Therefore, before any evaluation can be made as to the viability of the digit stream dual-task as the basis for a LE test (especially for listening conditions that resemble electric hearing), the issues with tracking need to be resolved.

5.8. Proposed solution

In line with the argument that it is not necessarily the individual tracking protocols that were failing per se, but rather the consecutive execution of more than one tracking, another experiment will now be pursued where only the channel tracking will be integrated within the procedure.

To ascertain an appropriate SNR (and attempt to equalise auditory performance), there will now be a choice of three fixed noise levels, which will individually be tested in order to find the SNR that elicits an auditory accuracy level of around 40-65% correct.

This range of accuracy is chosen in accordance with the reported finding that that LE measurement was optimal when the listener's auditory accuracy was around 50-70% correct (Gatehouse & Gordon, 1990). The accuracy range has been brought down marginally in order to accommodate the additional difficulty expected to be produced by the spectral degradation (i.e. caused by vocoding the auditory stimuli).

A new and larger cohort of NH controls will also be recruited to enable inferential statistical analysis.

5.9. Experiment 7

5.9.1. Hypotheses

With the amended protocol of only the channel tracking being implemented, and noise level manually chosen, the auditory performance achieved in quiet and in noise will be equalised. Accordingly, the following will be detected:

- LE will increase when spectral resolution decreases (i.e. visual accuracy will decrease when dual-tasking in quiet, relative to performing just the visual task);
- LE will increase even more when noise is introduced (i.e. visual accuracy will decrease when dual-tasking in noise, relative to dual-tasking in quiet).

These changes in visual accuracy will not be at the cost of resource allocation:

- Auditory accuracy will remain constant when the participant is dual-tasking (relative to that achieved when performing only the auditory task).
- Over-reporting will also be ameliorated.

In addition, there will be subjective corroboration of the behavioural data:

- Perceived effort will increase when spectral resolution decreases;
- Perceived effort will increase even more when noise is introduced.

5.9.2. Methods

5.9.2.1. Participants

30 adult normal hearing (NH) controls (14 females and 16 males, mean age: 44.0 years, \pm 15.1 S.D) were recruited. All participants had normal, or corrected-to-normal vision, and were native British English speakers.

5.9.2.2. Procedure

Location of testing: Testing sessions took place within the participants' home or office environments.

Primary task of the Digit Stream Dual-Task: Same as before.

Secondary task of the Digit Stream Dual-Task: Same as before.

Acoustic challenge: Same as before.

Vocoding parameters: Same as before.

Other LE indices: The same subjective ratings as before.

CI simulation training: The same as for Experiment 6 (i.e. a separate session totalling 90 minutes in duration) with only the following modifications:

- The removal of the training of vocoded stimuli at the participant's bespoke channel number in noise set at 7dB SNR and also at the participant's SRT. Instead, the noise level was set at 6dB SNR.
- The abandonment of noise tracking, with instead the individual testing of the primary task (of the Digit Stream Dual-Task) in three different noise levels: 6dB, 8dB and 10dB SNR (note, the auditory stimuli were vocoded at the participant's bespoke channel number as determined by channel tracking). The lowest SNR at which the participant's performance was still above 40% correct was used to define the participant's bespoke noise level.

For training conditions, see Figure 5.15. For overall protocol, see Table 5.4.

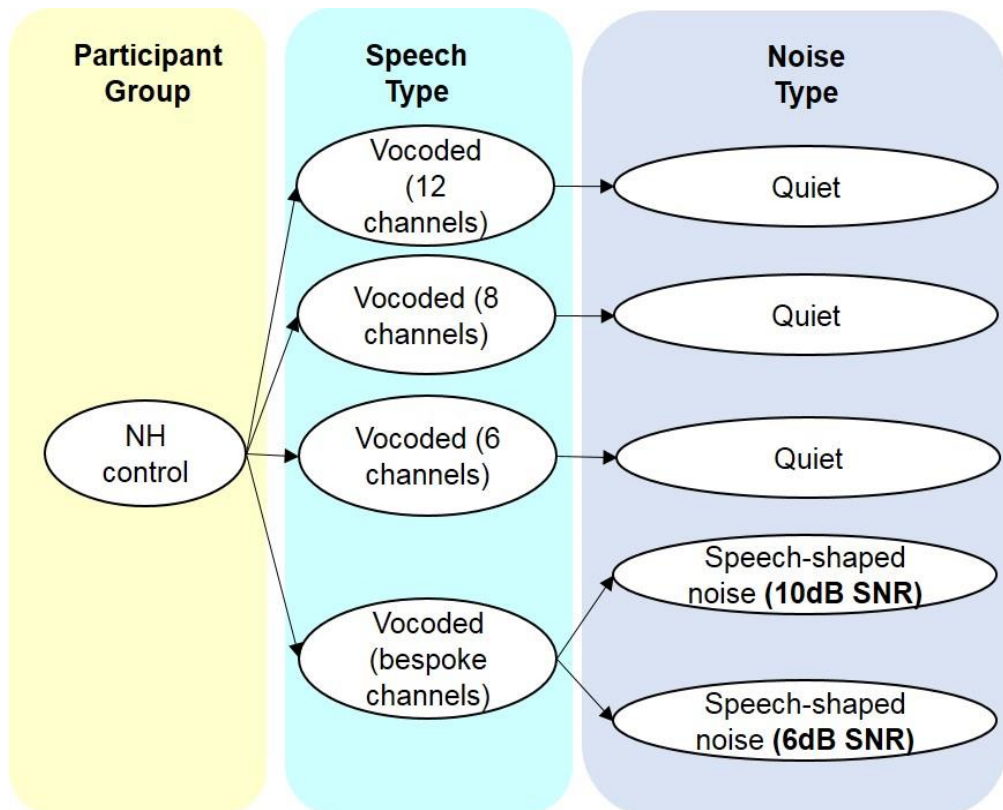


Figure 5.15: A schematic of conditions within CI simulation training (same structure for Figure 5.1).

5.9. Experiment 7

Table 5.4: Protocol for training in Experiment 7. Same detail provided regarding box shading as for Table 2.3. Please note no counter-balancing or randomisation apply in this protocol (i.e. all steps are executed in fixed order as listed).

Step	Training, Practice or Test?	Task Type	Speech Type	Noise Type	Total number of trials/ Total duration
1	Training	Indiana Jones	Vocoded (12 channels)	Quiet	2 minutes
2	Training	Indiana Jones	Vocoded (8 channels)	Quiet	5 minutes
3	Training	Indiana Jones	Vocoded (6 channels)	Quiet	5 minutes
4	Test	Adaptive tracking (tracking 75% correct)	Vocoded (variable channels)	Quiet	20 trials
5	Test	Adaptive tracking (tracking 75% correct)	Vocoded (variable channels)	Quiet	20 trials
6	Training	Indiana Jones	Vocoded (bespoke channels)	Speech-shaped noise at 10dB SNR	5 minutes
7	Training	Indiana Jones	Vocoded (bespoke channels)	Speech-shaped noise at 6dB SNR	10 minutes
8	Test	Auditory task only	Vocoded (bespoke channels)	Speech-shaped noise at 6dB SNR	12 trials
9	Test	Auditory task only	Vocoded (bespoke channels)	Speech-shaped noise at 8dB SNR	12 trials
10	Test	Auditory task only	Vocoded (bespoke channels)	Speech-shaped noise at 10dB SNR	12 trials
11	Practice	Visual task only at 25 digits	N/A	Quiet	10 trials
12	Practice	Dual at 25 digits	Vocoded (12 channels)	Quiet	12 trials

Adaptive tracking (channel only): The same protocol for channel tracking in Experiment 6 continued to be implemented, as well as the same method for deriving the participant's bespoke channel number from the two attempts of channel tracking.

Test conditions of Digit Stream Dual-Task (and trial structure): In total, there were six test conditions attempted (Figure 5.16), of which there were two conditions of speech type (i.e. vocoded auditory stimuli, set at 16 channels and also at the participant's bespoke number of channels), three conditions for task type (i.e. visual-only, auditory-only, or dual-task), as well as two conditions for noise type (i.e. quiet or noise with SNR set at the participant's bespoke level, which was either 6dB, 8dB or 10dB SNR). There were 18 trials for each test condition.

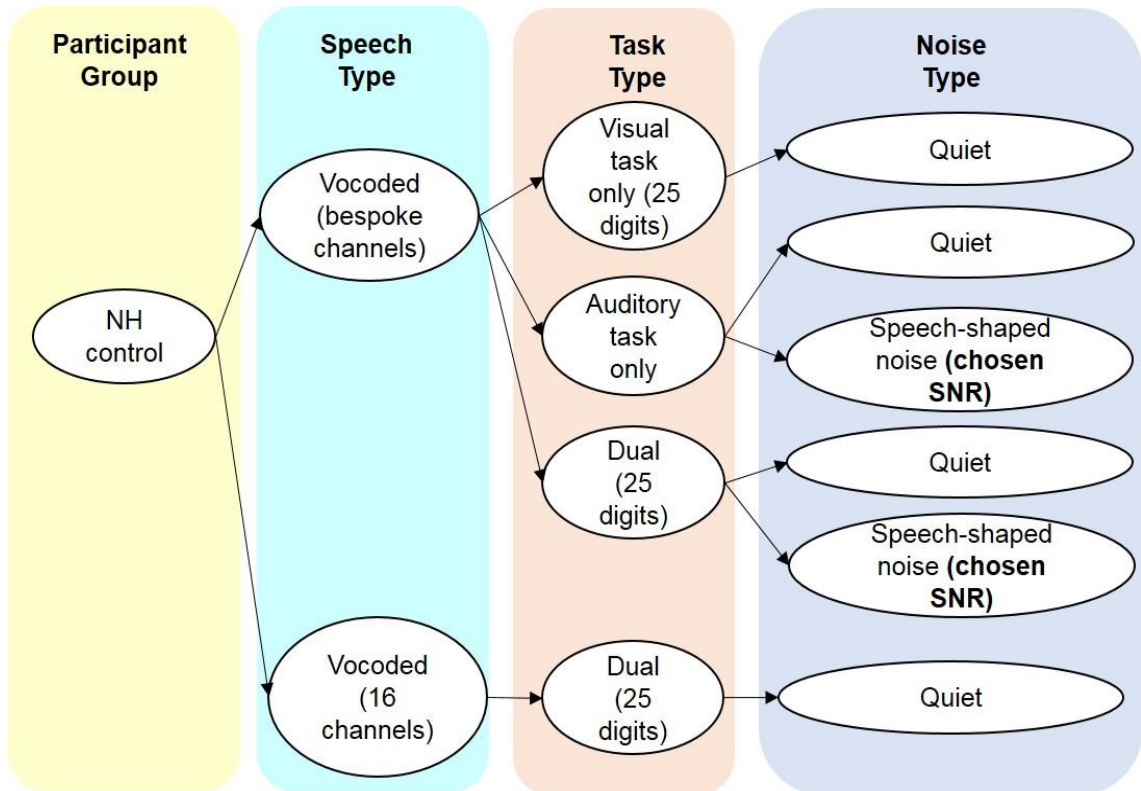


Figure 5.16: A schematic of test conditions (same structure as Figure 3.1).

Protocol: As in Experiment 6, the session for the testing of the Digit Stream Dual-Task was conducted after a compulsory minimum of 24 hours rest (which must involve at least one night of sleep). Overall protocol was the same as that utilised for Experiment 6 (see Table 5.5), except for the removal of a practice condition where the participant dual-tasked in quiet with spectral resolution set at 12 channels. All test conditions of the Digit Stream Dual-Task were counterbalanced and randomised across participants in an attempt to counteract training and/or learning effects (i.e. steps 5-10 in Table 5.5).

Statistical analysis: As for Experiment 4.

Table 5.5: Protocol for testing of the Digit Stream Dual-Task in Experiment 7. Same detail provided as for Table 2.3, including red numbers (with asterisks) indicating all the steps in the protocol where the order of the test conditions were randomised and counter-balanced together.

Step	Training, Practice or Test?	Task Type	Speech Type	Noise Type	Total number of trials/ Total duration
1	Training	Indiana Jones	Vocoded (bespoke channels)	Speech-shaped noise at chosen SNR	5 minutes
2	Practice	Visual task only at 25 digits	N/A	Quiet	10 trials
3	Practice	Dual at 25 digits	Vocoded (bespoke channels)	Quiet	18 trials
4	Practice	Dual at 25 digits	Vocoded (bespoke channels)	Speech-shaped noise at chosen SNR	18 trials
5 *	Test	Auditory task only	Vocoded (bespoke channels)	Quiet	18 trials
6 *	Test	Auditory task only	Vocoded (bespoke channels)	Speech-shaped noise at chosen SNR	18 trials
7 *	Test	Dual at 25 digits	Vocoded (bespoke channels)	Quiet	18 trials
8 *	Test	Dual at 25 digits	Vocoded (bespoke channels)	Speech-shaped noise at chosen SNR	18 trials
9 *	Test	Dual at 25 digits	Vocoded (16 channels)	Quiet	18 trials
10 *	Test	Visual task only at 25 digits	N/A	Quiet	18 trials

Power analysis:

Prior to recruitment of the participants, a power analysis was undertaken to gauge whether it was possible to recruit enough participants to achieve power for this experiment. This power calculation was executed by using the PASS 8 software (NCSS Statistical Software) to implement the two-sided McNemar test (for paired proportions: Schork & Williams, 1980, Machin et al., 1997). The significance level was set at 5%, and power at 80%.

The visual task outcomes of Experiment 4 formed the basis of the calculation (i.e. visual accuracy when dual-tasking at two different SNRs). To prevent the contamination caused by over-reporting, the target detection percentage for each trial was rendered binary, i.e. correct or incorrect. The sum of discordant proportions was then calculated (range: 0.4 to 0.55, $M = 0.74$). No time effect was also assumed (due to absence of effect of trial in Experiment 4).

This power analysis suggested that a sample size of over 60 participants was required to provide the attainable odds ratio of 3 (with a significance level of 0.05), indicating that effect size of the LE measure was likely to be small. Since time constraints prevented recruitment of this size, it is therefore likely that all inferential analyses pursued in this experiment will not achieve power.

5.9.3. Results and Discussion: Descriptive and inferential statistics

5.9.3.1. *LE measurement (secondary task performance): Visual accuracy*

Target detection tended to decrease as the test condition became systematically more difficult: i.e. with the requirement of dual-tasking, then with the reduction of spectral resolution, and finally with the addition of noise (Figure 5.17a). There was also considerable variability in visual accuracy achieved across all test conditions, even when the visual task was performed on its own (ranging from just below 60% to above 90%). In addition, the majority of NH controls obtained positive LE scores in both noise and in quiet, indicating that the acoustic challenge provided by the noise and also the vocoding increase the level of LE (as originally hypothesised: Figure 5.17b). However, several NH controls did exhibit negative LE scores in all listening conditions, which is contrary to expectation.

To ascertain the statistical significance of these trends, two models were generated. The first examined the degree of LE in quiet, by comparing visual accuracy when performing the visual task only, and when dual-tasking in quiet (at both levels of spectral resolution). This involved a mixed effects logistic regression model, with a single predictor of test condition. Participants were set as random effects (in the attempt to generalise the findings to the population). A statistically significant result ($F(2,87)=16.8$, $p<.001$) suggested that visual accuracy significantly differed across the three test conditions. Therefore, pairwise contrasts (Bonferroni-corrected) were pursued on the mean visual accuracy in each condition (Table 5.6). These comparisons revealed that LE significantly increased only when the spectral resolution was set at the participant's bespoke channel number ($p<.001$; for the easier spectral resolution of 16 channels: $p=.47$). Furthermore, when visual performances at both channel numbers were directly compared to each other, they significantly differed ($p<.001$), reflecting the increase in difficulty of spectral resolution (at bespoke level).

The second model investigated the level of LE exerted in noise, by comparing visual accuracy when dual-tasking in quiet (with spectral resolution set at the participant's bespoke channel number) and when dual-tasking in noise. Another mixed effects logistic regression model was applied, with a single predictor of test condition. Participants were again set as random effects. A statistically significant result ($F(1,58)=5.6$, $p=.021$) suggested that visual accuracy differed across the test conditions. This thus suggested that LE did increase when noise was added to the spectrally degraded auditory stimuli.

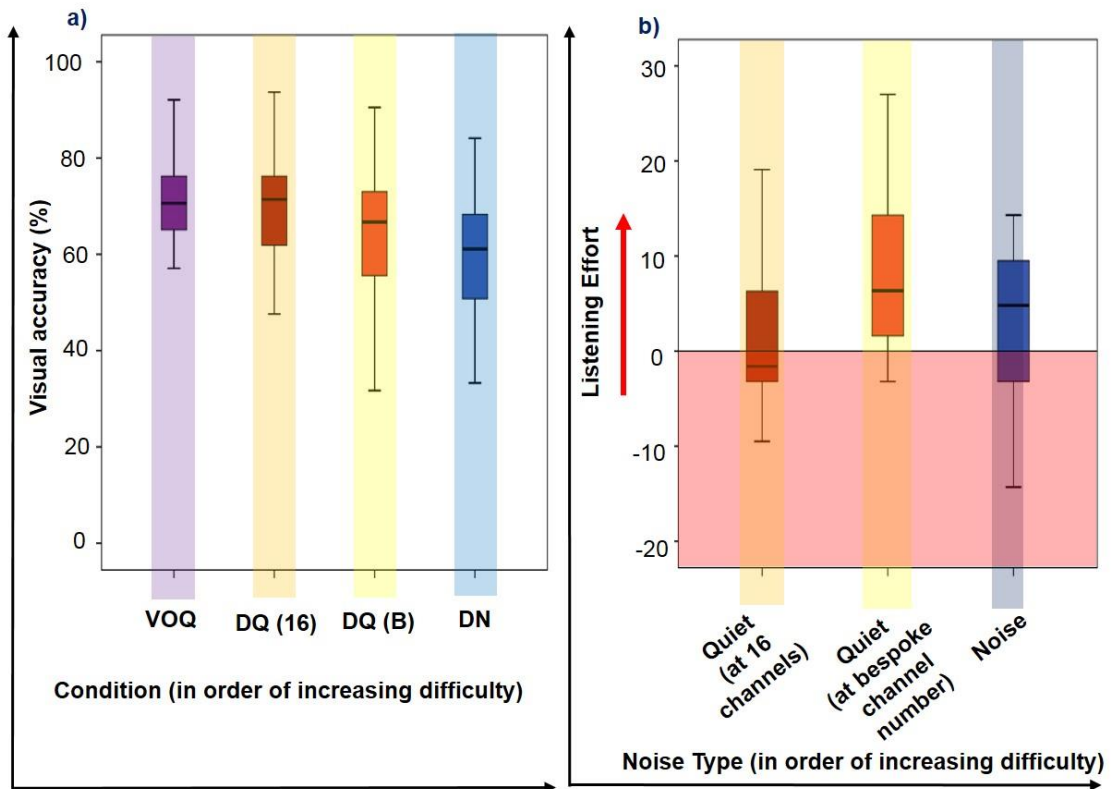


Figure 5.17: Boxplots of visual accuracy across test conditions.

VOQ = Visual task only in quiet, DQ (B) = Dual-tasking in quiet, with spectral resolution set at the participant's bespoke channel number (as obtained from the two attempts of channel tracking), DQ (16) = Dual-tasking in quiet, with spectral resolution set at 16 channels, DN = Dual-tasking in noise, with spectral resolution set at the participant's bespoke channel number and noise level set at the participant's bespoke SNR (either 6dB, 8dB or 10dB).

Red arrow indicates direction in which LE increases. Red shading indicates when LE has decreased in presence of challenge (i.e. either reduced spectral resolution or noise). Unit of LE = Difference in percentage.

Table 5.6: Mean and standard error of visual accuracy for each test condition

Test condition	Mean	Standard error
Visual only in quiet	71.3	1.6
Dual-tasking in quiet: 16 channels	70.3	1.9
Dual-tasking in quiet: bespoke channels	63.4	2.4
Dual-tasking in noise: bespoke channels & bespoke SNR	59.7	2.4

5.9.3.2. Evaluation of validity of LE measurement: Subjective ratings

To establish the validity of these LE scores, the subjective ratings were scrutinised to ascertain whether, or not, there was corroboration. It was found that, as the listening conditions became more challenging, ratings of perceived effort appeared to increase (Figure 5.18a). This was as hypothesised. However, the visual-only condition itself elicited variability so substantial that it encompassed the entire range of the rating scale (i.e. from “not at all hard work” all the way to the other extreme of “extremely hard work”).

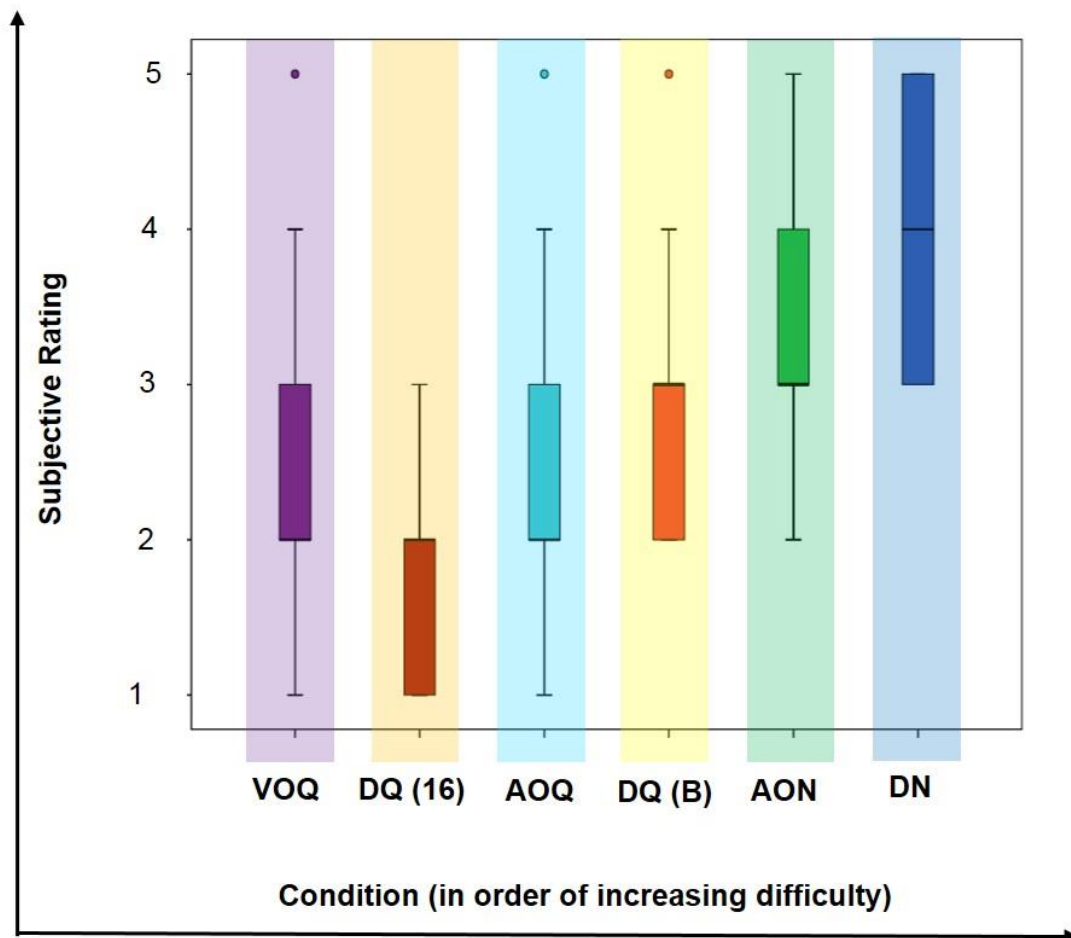


Figure 5.18: Boxplots of subjective ratings.

VOQ = Visual task only in quiet, DQ (B) = Dual-tasking in quiet, with spectral resolution set at the participant's bespoke channel number, DQ (16) = Dual-tasking in quiet, with spectral resolution set at 16 channels, DN = Dual-tasking in noise, with spectral resolution set at the participant's bespoke channel number and noise level set at the participant's bespoke SNR, AOQ = Auditory task only in quiet, AON = Auditory task only in noise set at the participant's bespoke SNR.

Subjective ratings: 1 = Not at all hard work, 2 = Quite hard work, 3 = Medium hard work, 4 = Very hard work, 5 = Extremely hard work.

To ascertain the statistical significance of the observed trend, a mixed effects linear model was generated with a single predictor of test condition (with six levels).

Participants were set as random effects. A statistically significant result ($F(5,145)=42.3$, $p<.001$) suggests that subjective ratings differed across test condition. Therefore, post-hoc pairwise comparisons (Bonferroni-corrected) were pursued on the mean subjective rating in each condition (Table 5.7).

Table 5.7: Mean and standard error of subjective ratings for each test condition

Test condition	Mean	Standard error
Visual only	2.2	0.2
Dual-tasking in quiet: 16 channels	1.8	0.1
Auditory only in quiet: bespoke channels	2.4	0.2
Dual-tasking at quiet: bespoke channels	2.9	0.1
Auditory only in noise	3.4	0.1
Dual-tasking in noise	4.0	0.2

Only two comparisons were found to be insignificant:

- Visual only versus dual-tasking in quiet at 16 channels ($p=.29$)
- Visual only versus auditory only in quiet at bespoke channels ($p=1$).

Otherwise, all comparisons were of significance with $p<.001$, except for the following:

- Visual only versus dual-tasking in quiet at bespoke channels ($p=.006$)
- Dual-tasking in quiet at 16 channels versus auditory only in quiet at bespoke channels ($p=.02$)
- Dual-tasking in quiet at bespoke channels versus auditory only in noise ($p=.04$)
- Auditory only in noise versus dual-tasking in noise ($p=.02$).

Therefore, there does appear to be subjective corroboration of visual performance.

However, before definitively concluding that LE measurement by the digit stream dual-task is reliable, there is a need to confirm primary task stability.

5.9.3.3. *Stability of primary task performance:*
Auditory accuracy

Overall, there did appear to be an effect of vocoding when spectral resolution was reduced to the participant's bespoke channel number (from that of 16 channels), as well as an effect of noise (Figure 5.19 a; for bespoke parameters applied to the auditory stimuli, see Figure 5.19 b). When primary task stability was calculated (Figure 5.19 c), there did appear to be deterioration in some NH controls when required to dual-task for both quiet and noise.

To ascertain the statistical significance of these trends, a 2x2 mixed effects logistic regression was applied to auditory accuracy percentage. Both predictors were categorical: task type (i.e. auditory-only or dual-task) and noise type (i.e. in quiet at bespoke channel number, or in noise). Interaction was allowed in this model, and participants were also set as random effects.

This model revealed an expected main effect of noise type: as noise was introduced, auditory accuracy significantly deteriorated ($F(1,116)=603.6$, $p<.001$). The main effect of task, however, was not significant ($F(1,116)=0.001$, $p=.98$), indicating a lack of difference between auditory-only and dual-tasking conditions. There was also no significant interaction ($F(1,116)=0.6$, $p=.46$). Thus, there appears to be stability in primary task performance (although, as explained earlier, there is the caveat that absence of effect is not necessarily evidence of absence).

5.9. Experiment 7

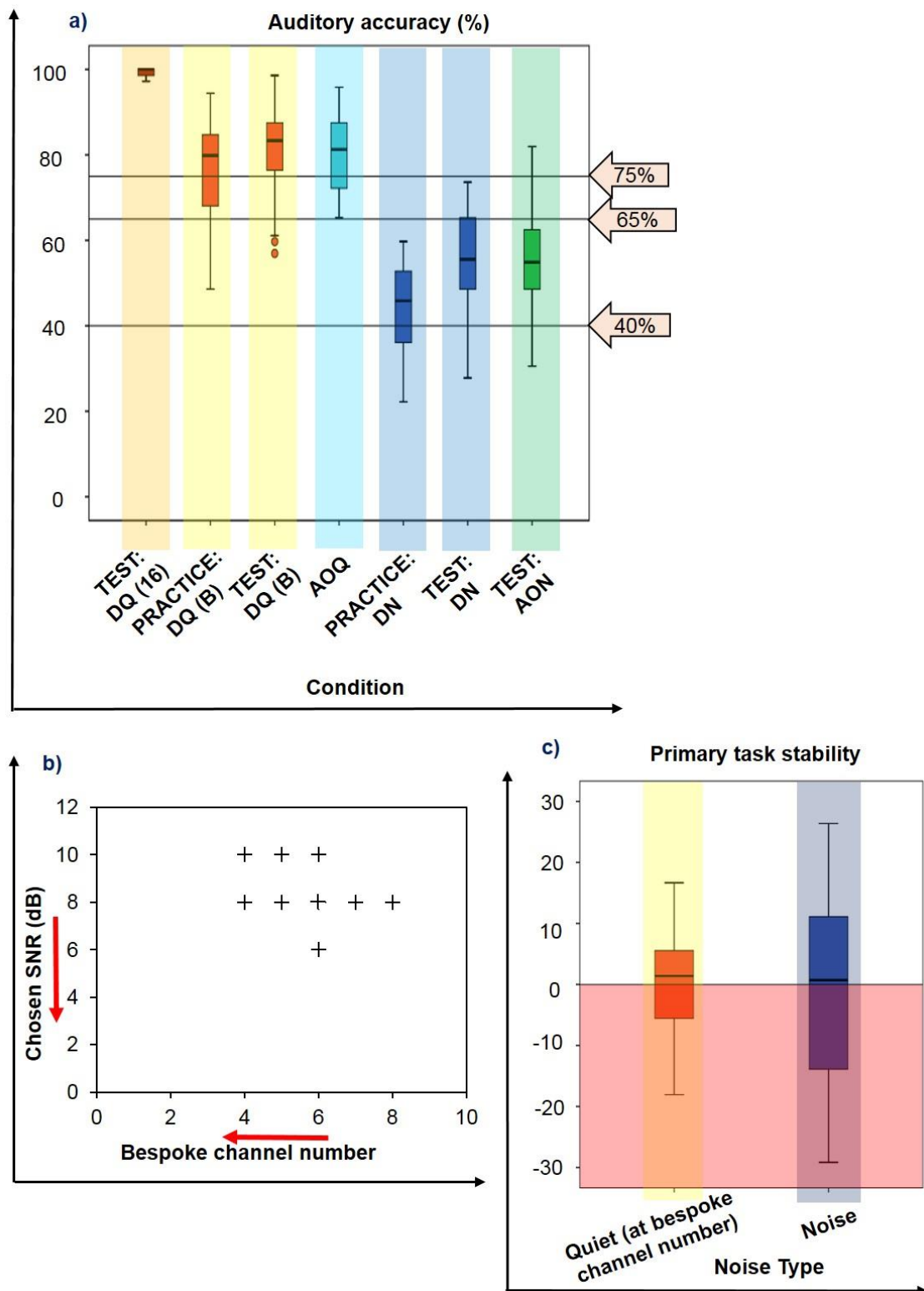


Figure 5.19: Boxplots of auditory accuracy (a) and of primary task stability (c), and scatterplot of testing parameters used (b).

DQ (B) = Dual-tasking in quiet, with spectral resolution set at the participant's bespoke channel number, DQ (16) = Dual-tasking in quiet, with spectral resolution set at 16 channels, DN = Dual-tasking in noise, with spectral resolution set at the participant's bespoke channel number and noise level set at the participant's bespoke SNR, AOQ = Auditory task only in quiet, AON = Auditory task only in noise at bespoke SNR. Red arrow indicates direction superiority of score (b). Red shading indicates when deterioration in auditory accuracy when dual-tasking (c). Unit of primary task stability = Difference in percentage.

5.9.3.4. *Evaluation of viability of LE measurement*

To ensure that LE measurement (i.e. secondary task performance) had not been compromised in any other way, the pattern of visual response was more closely scrutinised. As before, the number of trials for each type of visual response were tallied up for each target presentation frequency in each test condition (i.e. accurate response, reduced target detection, or over-reporting).

It emerged that, as the number of targets presented in the digit stream increased, the participants' visual response tended to drift towards that of reduced target detection (Figure 5.20). This was as expected. Unfortunately, it was also revealed that over-reporting was still occurring.

Therefore, the instances of over-reporting were isolated from the dataset for each test condition and target presentation frequency (Figure 5.21a). It was discovered that over-reporting occurred in the majority of the participants and within all test conditions. Not only this, there appeared to be the tendency that, as the test condition increased in difficulty, the incidence of over-reporting decreased (Figure 5.21b). In addition, as the number of targets presented in the digit stream increased, over-reporting incidence also decreased (Figure 5.21c).

To ascertain the statistical significance of the effect of test condition and target presentation frequency on over-reporting incidence, these variables were set as predictors for a mixed effects logistic regression model. Test condition was treated as a categorical variable (i.e. visual only, dual-tasking in quiet at the two levels of spectral resolution and dual-tasking in noise), whilst the number of targets presented in the digit stream was treated as continuous. Interaction was allowed, and participants were set as random effects.

This model revealed a significant main effect of test condition ($F(3,352)=3.6$, $p=.01$), as well as a significant main effect of target presentation frequency ($F(1,352)=53.7$, $p<.001$). Thus, over-reporting significantly decreased when SNR improved, and also when the number of targets increased. However, there was no significant interaction ($F(3,352)=1.5$, $p=.22$), indicating that the decrease in over-reporting with target number was similar for each SNR.

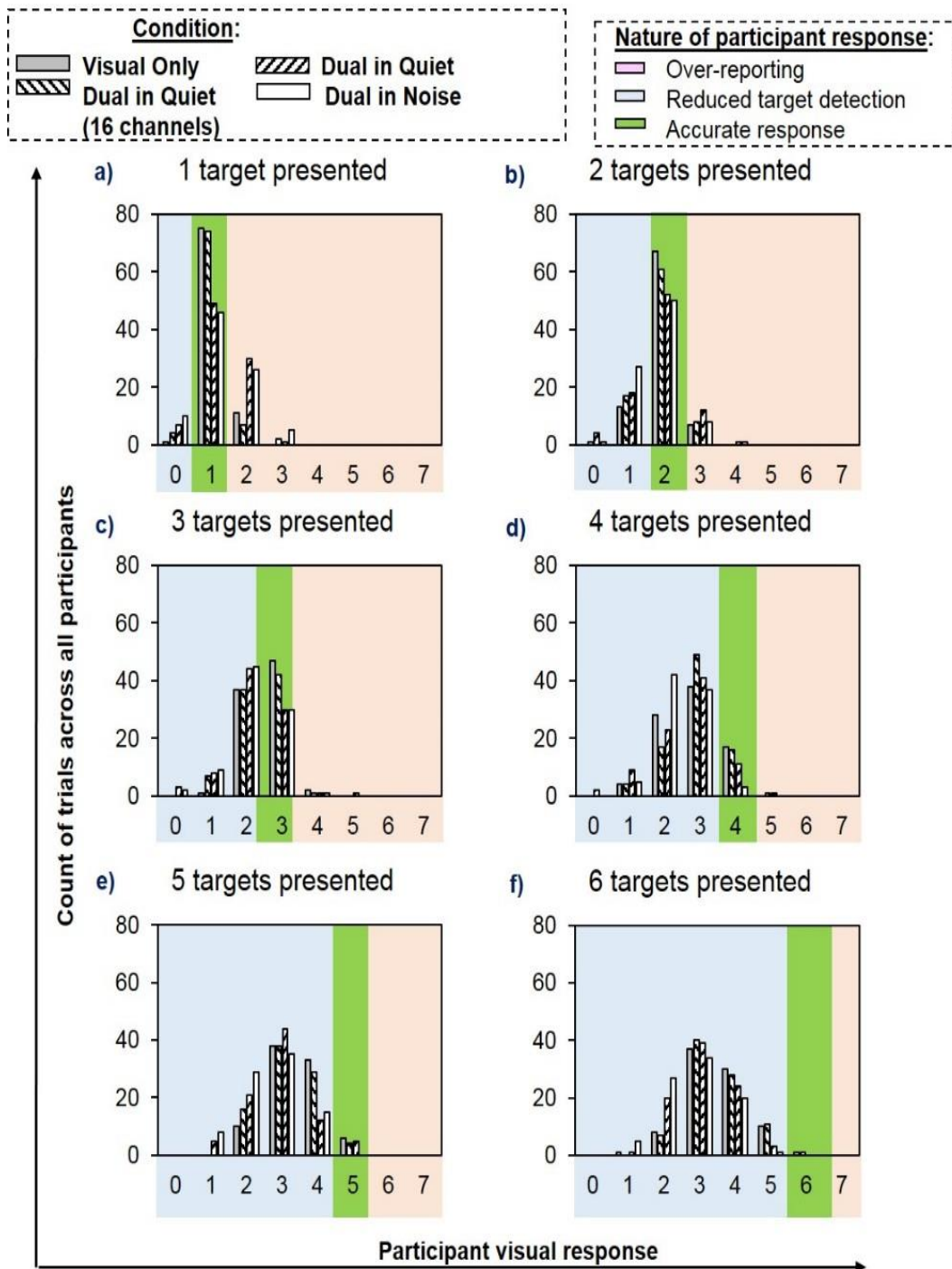


Figure 5.20: Histograms depicting the nature of participants' visual responses across trials (in each test condition) for each target presentation frequency

5.9. Experiment 7

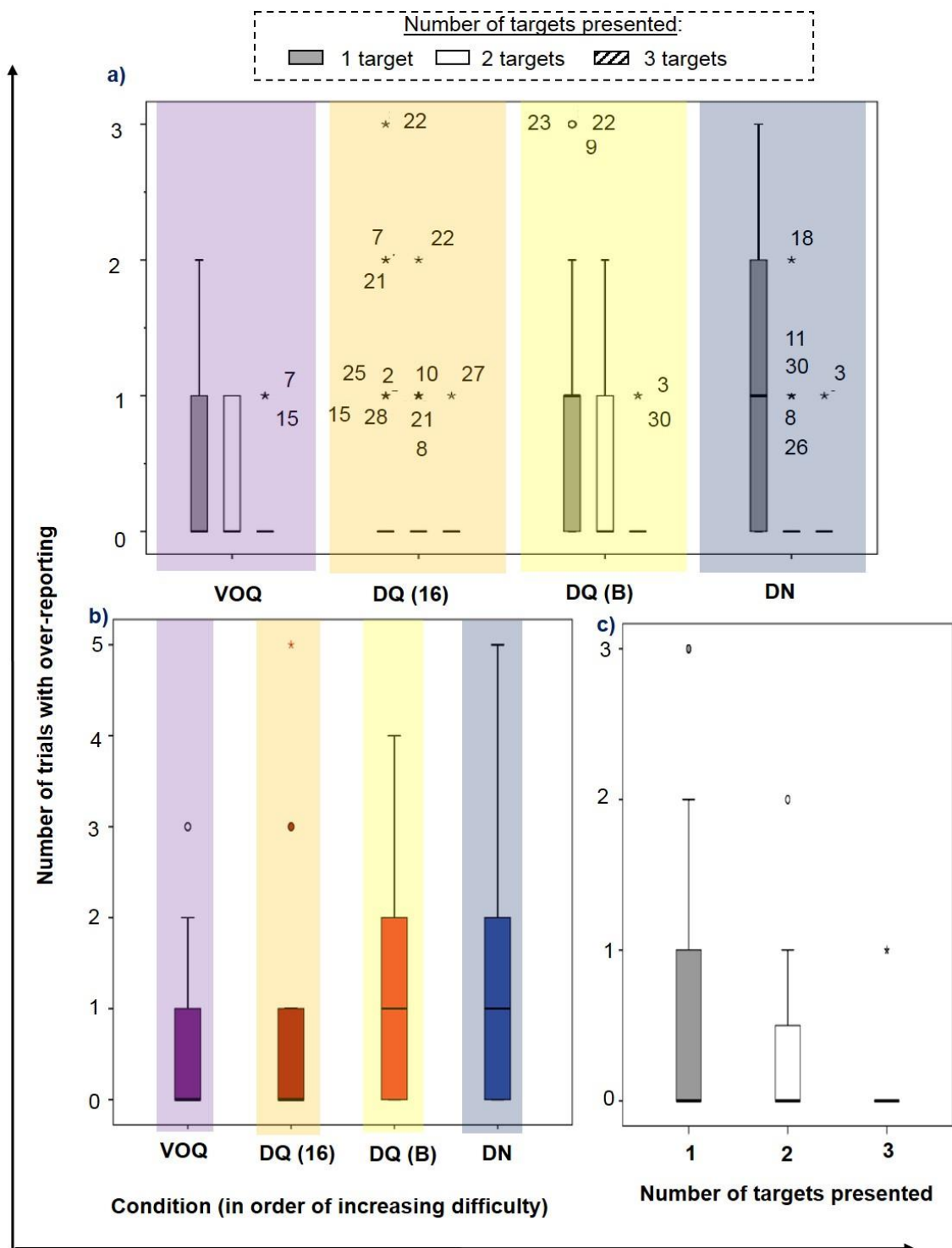


Figure 5.21: Boxplots of incidence of over-reporting, across all test conditions and target presentation frequency **(a)**, for each test condition (summated across all conditions of target presentation frequency: **b)**, and for each target presentation frequency (averaged across all test conditions: **c)**. Total number of trials for each test condition is 18.

VOQ = Visual task only in quiet, DQ (B) = Dual-tasking in quiet, with spectral resolution set at the participant's bespoke channel number, DQ (16) = Dual-tasking in quiet, with spectral resolution set at 16 channels; DN = Dual-tasking in noise, with spectral resolution set at the participant's bespoke channel number and noise level set at the participant's bespoke SNR. Asterisks (*) indicate that the occurrence of a particular value was an outlier. Superimposed numbers relate to the numerical identity of NH control(s) responsible for the given outlier value. Circles indicate extreme values.

Another aspect to confirm is the likelihood that auditory performance has been equalised across participants within both quiet and noisy listening conditions (i.e. whether the combination of adaptive tracking for spectral resolution and manual choice of noise level succeeded).

When test-retest reliability of the channel tracking (across the two tracking attempts) was assessed using the intraclass correlation coefficient (ICC), a relatively high level of agreement was found (ICC=0.60). This thus suggested a certain amount of consistency in auditory performance during the adaptive tracking.

Furthermore, when the auditory-only performance in quiet was plotted as a function of the participant's bespoke channel number (Figure 5.22 a), it was found that there was no significant correlation (Pearson's $r(30)=0.311$, $p=.095$). This lack of correlation was as desired, since this indicated a certain level of control of participant performance level.

In addition, none of the tracks and psychometric produced by the NH controls were found to be particularly aberrant (for examples of the two attempts of channel tracking, see Figure 5.22 c- f; for the tracking output for the remainder of NH controls, see Appendices 5.15-5.29; also, for the breakdown of how many participants required each category of channel number and SNR, see Appendix 30 a & b respectively).

Furthermore, when the auditory-only performance in noise was plotted as a function of chosen noise level (i.e. 6dB, 8dB, or 10dB SNR: Figure 5.22 b), the majority of NH controls' performances did lie between the required range of 40-65% correct. Importantly, there was also no significant correlation (Pearson's $r(30)=0.224$, $p=.234$). Thus, it seems that auditory performance had also been stabilised in noise. (For auditory data obtained during the trialling of the different noise levels compared to auditory accuracy achieved during the digit stream dual-task itself, see Figure 5.22 g).

The overall lack of correlation also has the potential implication that perceptual learning (or transfer) did not continue after training (i.e. during testing of the digit stream dual-task testing). Therefore, it seems that the CI simulations have now been successfully integrated within the digit stream dual-task.

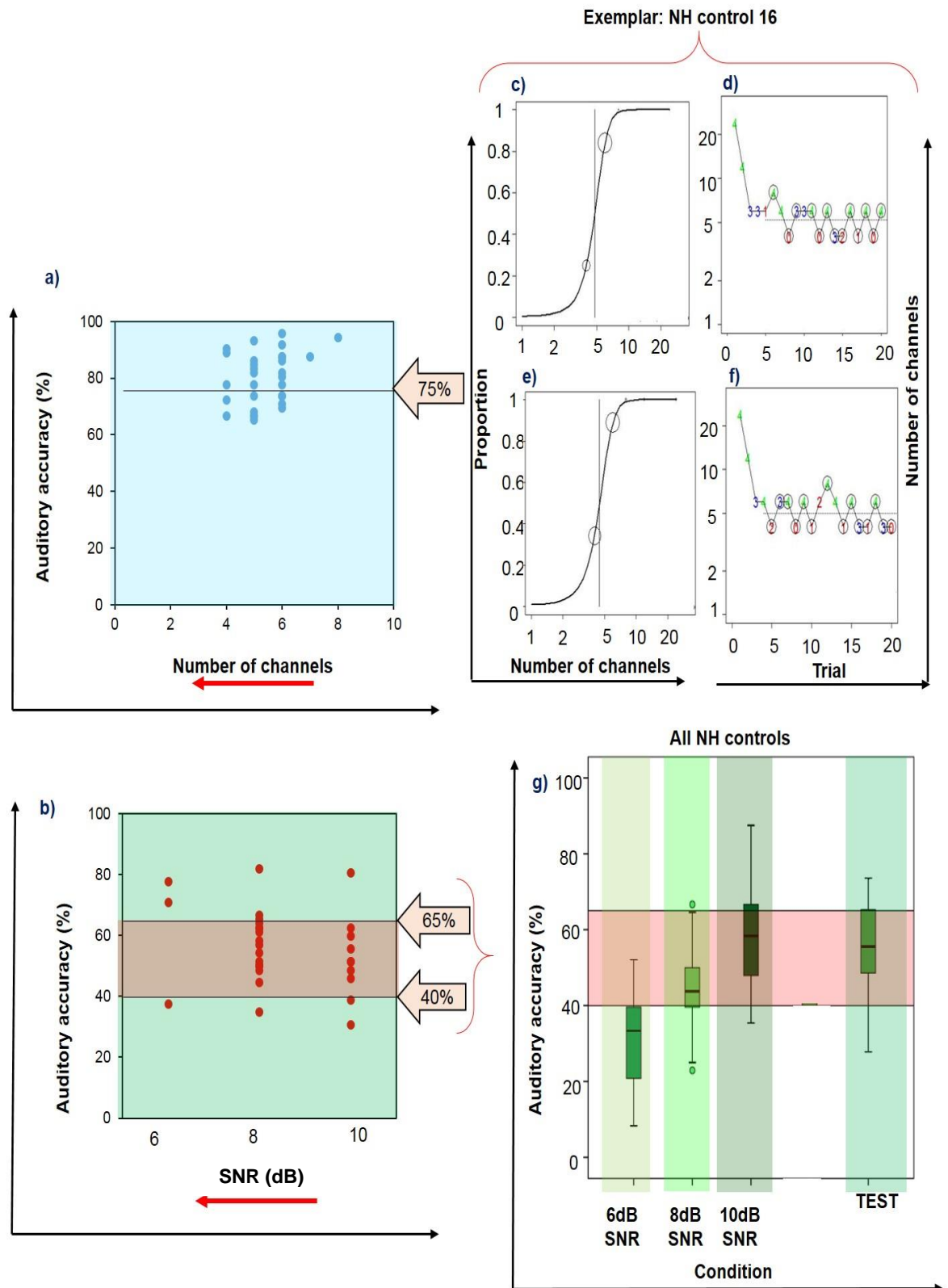


Figure 5.22: Scatterplots of auditory accuracy versus channel number and SNR (a, b), exemplars of psychometric functions and adaptive tracks for channel tracking (c, d, e, f), and boxplots of auditory performance in training and testing sessions (g). AOQ = Auditory task only in quiet, AON = Auditory task on its own in noise. Red arrow indicates superiority of score. Red shading indicates desired performance range for auditory accuracy in noise.

However, there remains a caveat to be considered: as effective as noise-excited vocoding techniques are in mimicking the spectral resolution and distortion of the CI input, there are still limitations in what they can reproduce with regards to the reality of electric hearing. After all, these CI simulations cannot recreate the physiological consequences of a deafened brain.

Indeed, as discussed within the literature review of Chapter 1, deafness can be conceptualised as a connectome disease, with collateral consequences in other neurocognitive systems beyond that of auditory processing (Kral et al., 2016; O'Donaghue et al., 2016). While CI simulations may emulate the disrupted auditory processing, they are unlikely to be able to induce the same extent of disturbance that is characteristic of a connectome disease.

5.10. Conclusions of Experiment 7

With the apparently successful incorporation of CI simulations, evaluation of the digit stream dual-task within a CI-related context is now possible. Also, with primary task stability potentially achieved, it seems likely that resource allocation was as designed during dual-task performance.

Interestingly, significant increases in LE were detected in quiet, as well as in noise. Thus, the acoustic challenge produced by the vocoding alone seems to already impose a cognitive toll, with this toll further increasing when noise is presented.

However, proof of concept (in terms of the digit stream dual-task's viability as a LE measure) can only be tentatively assumed. This is because the incidence of over-reporting has persisted and this may impact on the reliability of the measurement of visual accuracy. Also, it is necessary to recognise that CI simulations are unlikely to fully recreate the physiological consequences of deafness within the hearing brain.

Therefore, the next chapter will explore the trialling of the digit stream dual-task with twenty five adult CI users.

CHAPTER SIX:

Testing the refined digit stream dual-task with cochlear implant users (Experiment 8)

6.1. Introduction

As concluded by the previous chapter, the feasibility of the digit stream dual-task as a form of LE measurement for CI users is only possible if the following caveats are addressed: firstly, CI simulations are limited in their ability to recreate electric hearing; and secondly, the continued over-reporting may compromise reliability of LE measurement.

Therefore, the final experiment of this thesis will trial the digit stream dual-task in actual CI users and continue to monitor and analyse over-reporting.

6.2. Required modifications for CI testing

The digit stream dual-task will be tested within both quiet and noisy listening conditions. However, heterogeneity within the CI users' ability to successfully perceive speech in quiet is predicted (Blamey et al., 2013; Anderson et al., 2017; Holden et al., 2013; Lazard et al., 2010). This then has implications for the adaptive tracking being used to define the noise level applied (and, therefore, match auditory performance achieved across participants). More specifically, there are consequences related to the tracking criterion.

All adaptive tracking thus far (within this thesis) has utilised a fixed tracking criterion (e.g. ascertaining the participant's SRT). However, this is problematic for CI testing because, if the CI users are already showing wide variations in their ability to hear in quiet, the fixed tracking criterion is then unlikely to be able to equalise auditory performance in response to the noise across the CI cohort. This is because the CI users may differ in how easily they can achieve the defined performance level in noise, according to what extent they are already struggling in quiet.

A potential solution to accommodate this potential variability in the starting point of a given CI user's ability in speech perception (in quiet) is to tailor the tracking criterion used for each CI user. Accordingly, the definition for this bespoke tracking criterion will be the following: the SNR at which the CI users' performance in noise is at 50% of the accuracy of their performance in quiet (e.g. should a participant be achieving 80% accuracy in quiet, then the tracking criterion will find the SNR that induces 40% accuracy in noise).

6.3. Experiment 8

6.3.1. Hypotheses

The use of bespoke tracking criteria will succeed in equalising the auditory performance achieved in noise.

Therefore, when noise is introduced into the digit stream dual-task, the following will occur:

- Visual accuracy will decrease (i.e. LE will increase).

This change in visual accuracy will be corroborated by the subjective ratings, i.e.:

- Perceived effort will increase as LE scores increase.

In addition, this change in visual accuracy will not be at the cost of resource allocation:

- Auditory accuracy will remain constant when the participant is dual-tasking (relative to that achieved when performing only the auditory task).
- Over-reporting will also be ameliorated.

6.3.2. Methods

6.3.2.1. Participants

25 adult CI users (15 females and 10 males) were recruited. All participants had normal, or corrected-to-normal vision, and were native British English speakers. The demographics (Figure 6.1) reflect recruitment of a range of CI users, both those who have been established users since the introduction of the paediatric implantation programme in the UK (i.e. from 1990 onwards), to those aged between 30 and 67 years, who have only been able to access the CI programmes more recently (i.e. within the last five years). The mean age of this cohort was 33.7 years (± 13.67 S.D), with the mean CI experience being 13.7 years (± 7.94 S.D). The majority of these participants were unilateral implantees, and were pre-lingually deafened (mean duration of deafness: 29.3 years, ± 12.21 S.D).

Also, although most individuals had a known medical cause for their deafness, a substantial proportion of the participants had unknown aetiology for their disability. The majority of the cohort preferred oral communication, but did possess knowledge of sign language (either British Sign Language or Sign-Supported English). All participants who volunteered for testing proved to be successful CI users, as evidenced by many possessing higher level qualifications (including postgraduate degrees).

6.3. Experiment 8

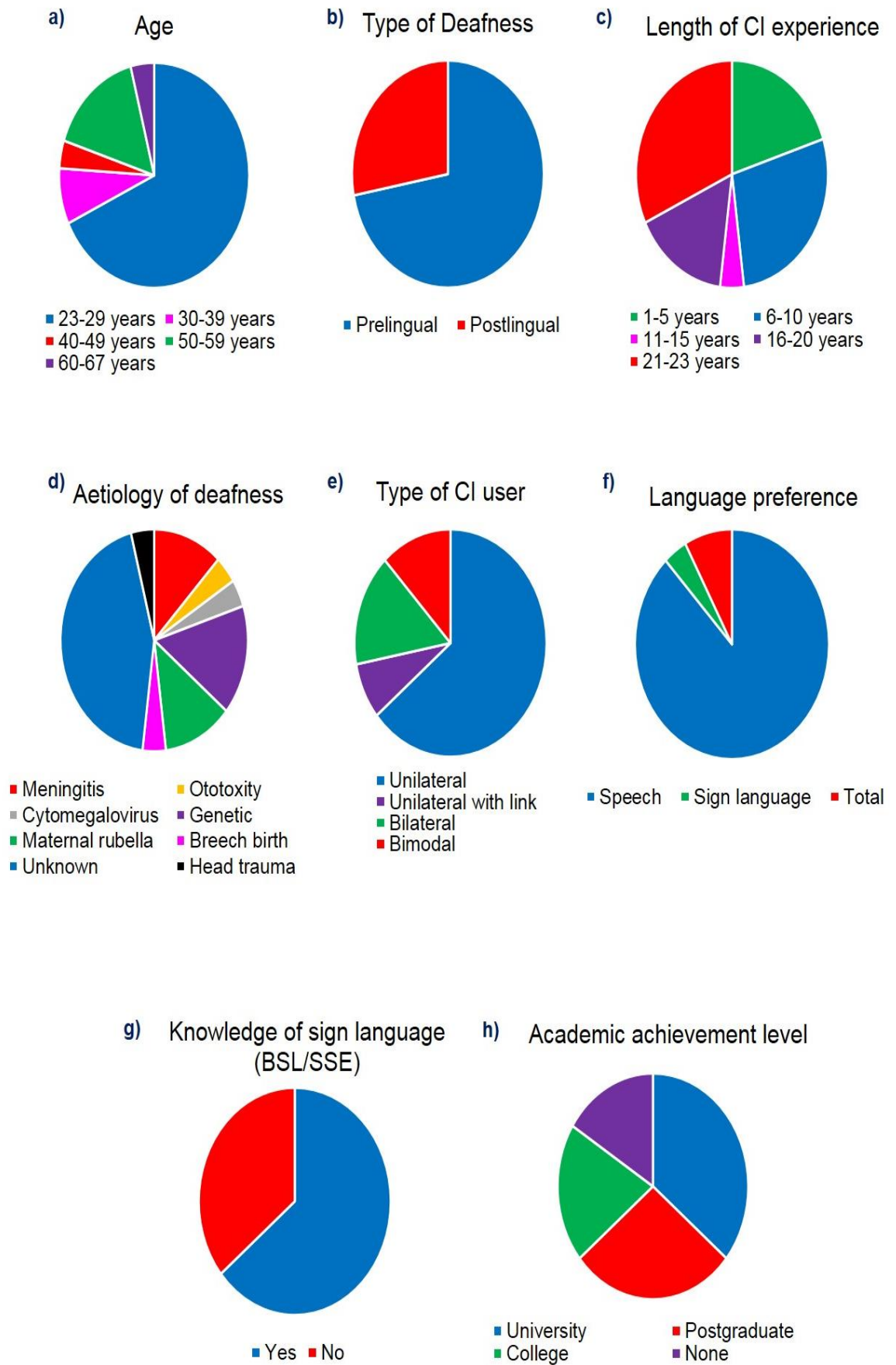


Figure 6.1: Pie charts representing the demographics of the CI users recruited.

6.3.2.2. Procedure

Location of testing: Testing sessions took place within the participants' home or office environment.

Primary task of the Digit Stream Dual-Task: Same as before, except that all auditory stimuli (for both tracking and testing) were delivered free-field via a calibrated MS101II monitor speaker (Yamaha). Overall sound pressure level remained at 70dB SPL at all times (including conditions with noise).

Secondary task of the Digit Stream Dual-Task: Same as before.

Acoustic challenge: Instead of speech-shaped noise, a multi-talker babble (comprised of 20 talkers) was utilised.

The rationale for this change was the recognition that CI users are generally more susceptible to background noise than NH controls listening to comparable CI simulations, particularly when the noise is dynamic, or of competing speech (Friesen et al., 2001; Fu & Nogaki, 2005; Nelson et al., 2003). Therefore, in an attempt to generate LE levels that might occur in the real-world scenario, a masker that more closely resembles what the CI user might encounter in everyday life (i.e. the multi-talker babble) was utilised in this experiment. It was hoped that, by introducing a more ecologically valid acoustic challenge, LE measurement in this experiment would become more sensitive, bearing in mind the additional constraints of CI hearing.

Adaptive tracking: The same software and stimuli as used for Experiment 6 was utilised. The same protocol was also used, but with the following changes:

- The tracking criterion was now defined to be the SNR eliciting 50% accuracy of the participant's performance in quiet.
- The participant performed the primary task (of the Digit Stream Dual-Task) twice in quiet before undergoing tracking (the average accuracy was calculated to derive the auditory performance in quiet that was part of defining the tracking criterion).

This, therefore, had implications for the adaptive algorithm used. The algorithm now implemented was based on the weighted up-down method developed by Kaernbach (1991). This algorithm was essentially based on the premise that each correct response led to a decrease in SNR on the next trial, whilst every correct response led to an increase in SNR. The key difference with this algorithm, however, was that the step size (S) used for upward (*up*) versus downward (*down*) changes was now tailored according to the following equation: $S_{up} P = S_{down} (1-P)$, where P is the convergence point (i.e. the desired performance level).

This adaptive tracking technique was chosen because it allows the tracking of an arbitrary performance level, therefore enabling the implementation of a bespoke tracking criterion.

The starting SNR for all CI users was 20dB (consistent with previous tracking). This also set the maximum SNR allowed in the adaptive tracking. The step size was initially 10dB until the first reversal (i.e. when their response goes in the opposite direction to the previous response), after which the adaptive algorithm began following Kaernbach's formula. However, this initial step size of 10dB only occurred if the participant achieved a correct response within the first three trials. A correct response was defined to be more than 2 key words (of the BEL sentence) correct. If not enough key words were detected within the first three trials, then the Kaernbach's formula immediately applied to the step size for any change in SNR.

If the participant achieved fewer than 2 key words, this was classified as an incorrect answer. However, if the participant achieved comprehension of 2 key words within the BEL sentence, there was no change in SNR for the next trial.

A minimum of 35% was set for the tracking criterion. This was due to the concern that, if the tracking criterion was set any lower, the performance level would be in such close proximity to (if not already at) floor level that the tracking protocol would be unable to derive any meaningful data.

The participant's bespoke SNR value was calculated by taking the mean of all levels visited (in SNR) at the final even number of reversals (if any reversals had been achieved). Psychometric functions also were computed (consistent with other experiments). However, should it be the case that the participant never achieved more than 2 key words correct with the SNR of 20dB, the SNR never increased for the entirety of the tracking.

As there were two attempts for the noise tracking, should better performance be obtained on the second run, the SNR value from this second run was utilised. If the best performance was achieved on the first run, an average of the values obtained in the two attempts were taken to comprise the participant's bespoke SNR. A second noise level was also derived from this bespoke SNR. This was intended to be an easier noise condition, so the SNR was increased by 5dB (on top of the bespoke level).

However, if the SNR had never decreased from 20dB during either of the tracking attempts, the participant's bespoke SNR was set at 20dB. Since 20dB denoted the maximum SNR allowed, this had the consequence that a second noise condition could not be executed with these participants.

Other LE indices: The same subjective ratings as before.

Test conditions of the Digit Stream Dual-Task (and trial structure): In total, there were seven test conditions attempted (Figure 6.2), of which there was one speech type (i.e. normal auditory stimuli), three conditions for task type (i.e. visual only, auditory-only, or dual-task), as well as three conditions for noise type (i.e. quiet, noise at the bespoke SNR, and noise at 5dB above the bespoke SNR).

For five of these test conditions, the participants were required to perform them twice (the only two conditions that did not have a second run were the auditory-only test conditions involving noise). These two runs were in order to gauge the potential test-retest reliability of the Digit Stream Dual-Task. There were 18 trials for each test condition (and each run).

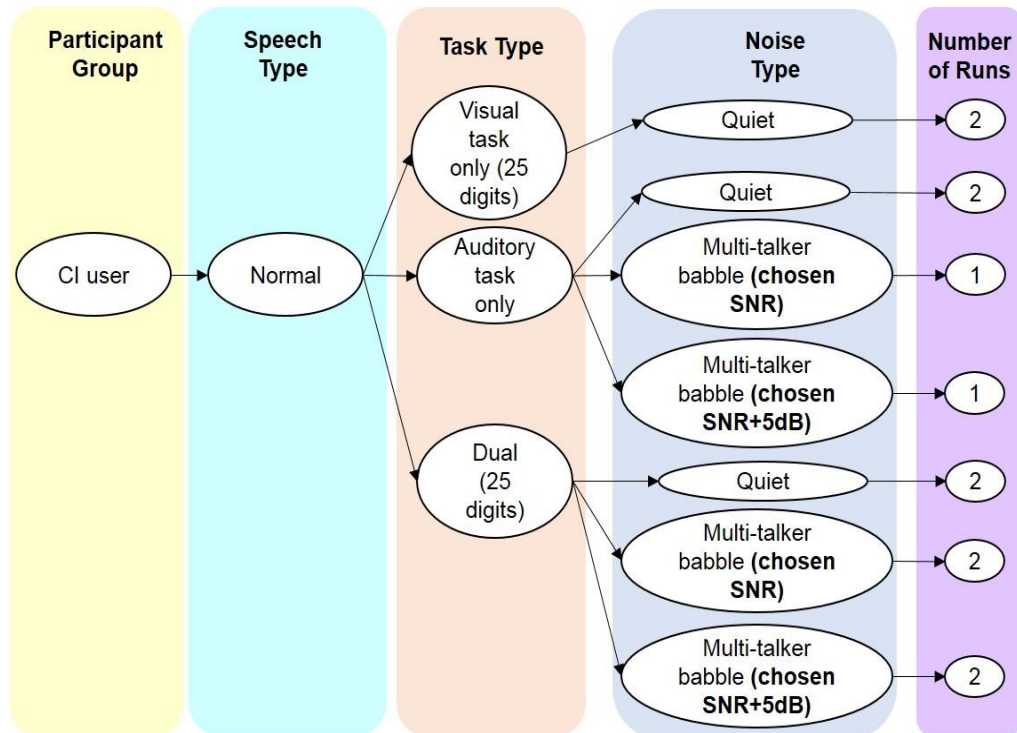


Figure 6.2: A schematic of test conditions (same structure as Figure 3.1 with the addition of how many runs applied for each test condition).

Protocol: The testing session amounted to 2 hours in total. For overall protocol, see Table 6.1. The order of test conditions within the block of first runs (steps 9-13) and also within block of second runs (steps 14-16) were counter-balanced, with the participant always finishing with the visual-only task. The session began with a practice run where CI users were provided 5 trials of the auditory task on its own in quiet, as well as 10 trials of the visual task on its own (also presented in quiet) and 10 trials for the dual-task condition (which was in quiet as well).

As aforementioned, participants whose bespoke SNR was 20dB could not undergo testing within the auditory-only or dual-tasking conditions involving the noise level of 5dB above the bespoke SNR.

Statistical analysis: As for Experiment 7, with the exception of the power calculation.

The rationale underlying this decision was that the previous power calculation (that was executed on NH controls in normal listening conditions) had already identified that the sample size needed to be 60 or more participants. In light of the anticipated increase in individual variation caused by the clinical heterogeneity of the CI population, there was a high likelihood that the sample size required to achieve power for CI users would become even larger (rendering recruitment of a sufficient size unachievable for this study).

Nevertheless, despite the predicted lack of power, this experiment was still undertaken because there is a need to generate baseline data about LE in the CI population. Once this baseline data has been gleaned, it may then be possible to tailor experimental design to produce studies capable of achieving power within this clinical population.

6.3. Experiment 8

Table 6.1: Protocol for Experiment 8. Same detail provided as for Table 2.3, including red numbers (with asterisks) indicating all the steps in the protocol where the order of the test conditions were randomised and counter-balanced together.

Step	Training, Practice or Test?	Task Type	Speech Type	Noise Type	Total number of trials/ Total duration
1	<i>Practice</i>	Auditory task only	Normal	Quiet	5 trials
2	<i>Practice</i>	Visual task only at 25 digits	N/A	Quiet	10 trials
3	<i>Practice</i>	Dual at 25 digits	Normal	Quiet	10 trials
4	Test	Visual task only at 25 digits (first run)	N/A	Quiet	18 trials
5	Test	Auditory task only (first run)	Normal	Quiet	18 trials
6	Test	Auditory task only (second run)	Normal	Quiet	18 trials
7	Test	Adaptive tracking (tracking half of percentage accuracy in quiet)	Normal	Multi-talker babble at variable SNR	20 trials
8	Test	Adaptive tracking (tracking half of percentage accuracy in quiet)	Normal	Multi-talker babble at variable SNR	20 trials
9 *	Test	Auditory task only (first run)	Normal	Multi-talker babble at chosen SNR	18 trials
10 *	Test	Auditory task only (first run)	Normal	Multi-talker babble at chosen SNR +5dB	18 trials
11 *	Test	Dual at 25 digits (first run)	Normal	Quiet	18 trials
12 *	Test	Dual at 25 digits (first run)	Normal	Multi-talker babble at chosen SNR	18 trials
13 *	Test	Dual at 25 digits (first run)	Normal	Multi-talker babble at chosen SNR +5dB	18 trials
14 *	Test	Dual at 25 digits (second run)	Normal	Quiet	18 trials
15 *	Test	Dual at 25 digits (second run)	Normal	Multi-talker babble at chosen SNR	18 trials
16 *	Test	Dual at 25 digits (second run)	Normal	Multi-talker babble at chosen SNR +5dB	18 trials
17	Test	Visual task only at 25 digits (second run)	N/A	Quiet	18 trials

6.3.3. Results and Discussion: Descriptive and inferential statistics

6.3.3.1. *LE measurement (secondary task performance): Visual accuracy*

An average was taken of the two runs of each test condition (see Appendix 6.1 for the visual accuracy scores obtained in the two attempts for each condition).

Overall, there was the tendency for the average target detection to decline when dual-tasking, with even further decline when noise was present (Figure 6.3a). There was also considerable variability in visual accuracy achieved across all test conditions, even with the visual task on its own (ranging from 50% to 90%). In addition, the majority of CI users obtained positive LE scores in both noise levels (as originally hypothesised: Figure 6.3b). Interestingly, the CI users also obtained positive LE scores in quiet, which was not predicted. However, several CI users did exhibit negative LE scores in all listening conditions. Thus, to ascertain the statistical significance of these trends, two models were generated.

The first examined the degree of LE in quiet, by comparing visual accuracy when performing the visual task only, and when dual-tasking in quiet. This involved a mixed effects logistic regression model, with a single predictor of test condition. Participants were set as random effects (in the attempt to generalise the findings to the population). A statistically significant result ($F(1,48)=22.2$, $p<.001$) suggested that visual accuracy significantly deteriorated when dual-tasking in quiet. Thus, it seems that LE in quiet was indeed sizeable.

The second model investigated the level of LE exerted in noise, by comparing visual accuracy when dual-tasking in quiet and when dual-tasking in noise at both noise levels (i.e. the participant's bespoke SNR and 5dB above this SNR). Another mixed effects logistic regression model was applied, with a single predictor of test condition. Participants were again set as random effects. The outcome was nearly significant ($F(2, 65)=3.01$, $p=.056$), suggesting that visual accuracy did not differ substantially across test condition. However, pairwise contrasts (Bonferroni-corrected) were pursued to more closely interrogate mean visual accuracy in each test condition (Table 6.2). These comparisons revealed that the increase in LE was of borderline significance ($p=.05$) when the noise level was set at the participant's bespoke SNR. At the easier SNR, no significant changes in visual accuracy were detected ($p=.40$). Furthermore, visual accuracy did not significantly differ between dual-tasking at the two noise levels ($p=.40$).

6.3. Experiment 8

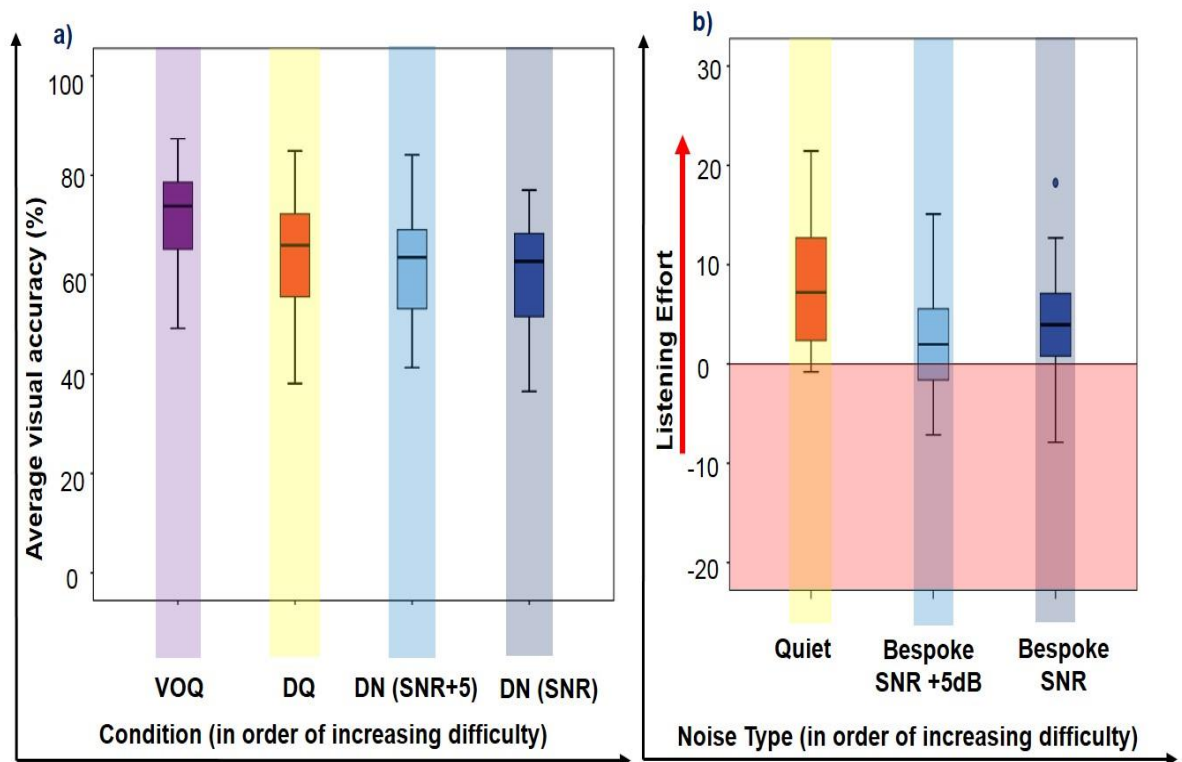


Figure 6.3: Boxplots of visual accuracy (a) and LE (b). Unit of LE = Difference in percentage.

DQ = Dual-tasking in quiet, DN (SNR) = Dual-tasking in noise at bespoke SNR, DN (SNR+5) = Dual-tasking in noise at 5dB above the participant's bespoke SNR.

Red arrow indicates direction in which LE increases. Red shading indicates when LE has decreased in presence of acoustic challenge.

Table 6.2: Mean and standard error of visual accuracy for each test condition

Test condition	Mean	Standard error
Visual only in quiet	72.4	2.0
Dual-tasking in quiet	64.7	2.3
Dual-tasking in noise at 5dB above bespoke SNR	62.3	2.9
Dual-tasking in noise at bespoke SNR	60.6	2.3

One possible interpretation of the lack of statistical significance in noise is that there genuinely may have been no changes in LE level. This is a tenable notion if it was indeed the case that listening in quiet did require considerable LE. This could then have the consequence of little margin being left for any further increase in LE when noise was introduced.

To further interrogate the nature of LE experienced by the CI users, subjective ratings warrant closer examination.

6.3.3.2. *Evaluation of validity of LE measurement:
Subjective ratings*

An average was taken of the two runs for each test condition (see Appendix 6.2 for the subjective ratings obtained in the two attempts for each applicable condition). As the test condition became more difficult, especially with the introduction of noise, perceived effort shifted towards that of “Extremely hard work” (Figure 6.4). It was notable that, in all test conditions, there was always at least one CI user who rated perceived effort as “Extremely hard work”, irrespective of whether there was noise (or not), or whether they were dual-tasking (or not).

To ascertain the statistical significance of the observed trend, a mixed effects linear model was generated with a single predictor of test condition (with seven levels). Participants were set as random effects. A statistically significant result ($F(6,131)=18.2$, $p<.001$) suggests that subjective ratings differed across test condition. Therefore, post-hoc pairwise comparisons (Bonferroni-corrected) were pursued on the mean subjective rating in each condition (Table 6.3).

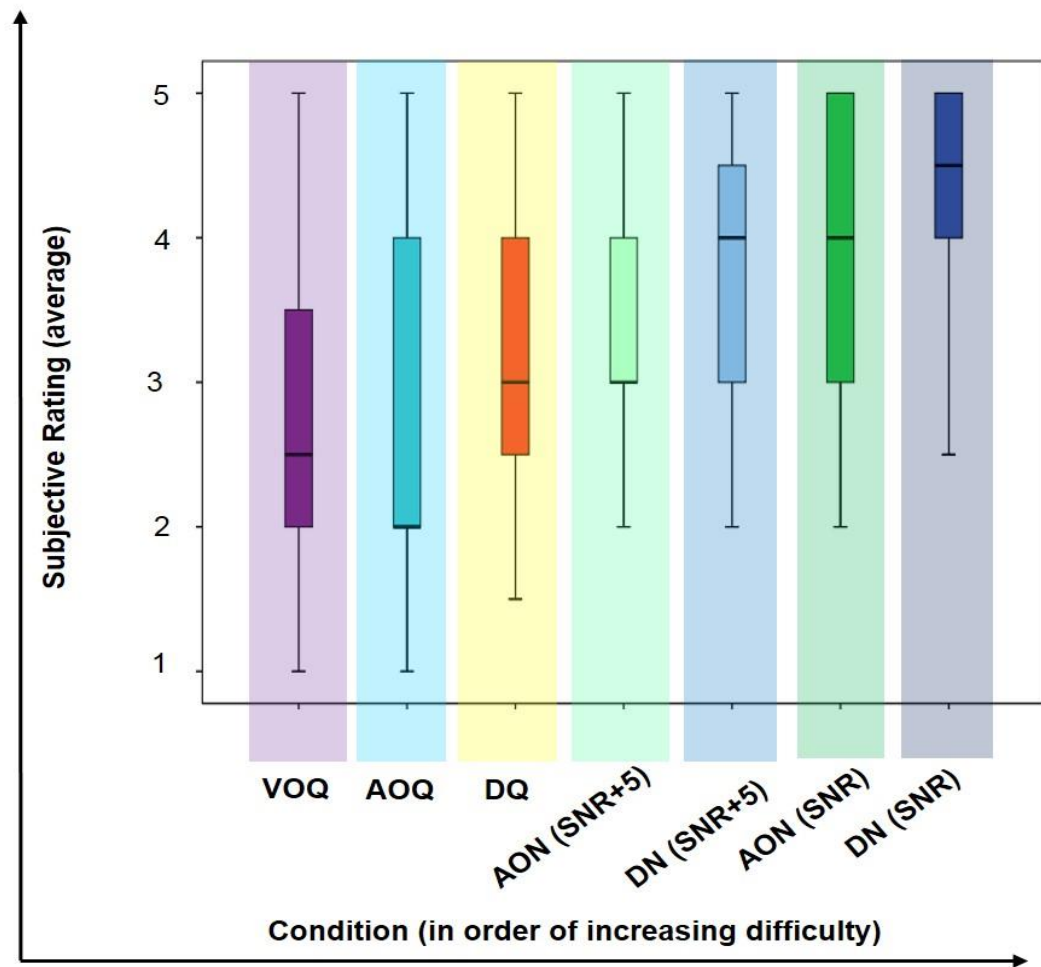


Figure 6.4: Boxplots of subjective ratings across test condition.

VOQ = Visual task only in quiet, AOQ = Auditory task only in quiet, DQ = Dual-tasking in quiet, AON (SNR) = Auditory task only in noise at bespoke SNR, DN (SNR) = Dual-tasking in noise at bespoke SNR, AON (SNR+5) = Auditory task only in noise at 5dB above bespoke SNR, DN (SNR+5) = Dual-tasking in noise at 5dB above bespoke SNR.

Subjective ratings: 1= Not at all hard work, 2 = Quite hard work, 3 = Medium hard work, 4 = Very hard work, 5 = Extremely hard work.

Table 6.3: Mean and standard error of subjective ratings for each test condition

Test condition	Mean	Standard error
Visual only	2.7	0.2
Auditory only in quiet	2.6	0.2
Dual-tasking in quiet	3.2	0.2
Auditory only in noise at 5dB above bespoke SNR	3.5	0.2
Dual-tasking in noise at 5dB above bespoke SNR	3.9	0.2
Auditory only in noise at bespoke SNR	3.9	0.2
Dual-tasking in noise at bespoke SNR	4.4	0.1

It was revealed that over half of these pairwise comparisons reached significance (Table 6.4). However, interestingly, there were many comparisons that did not achieve statistical significance (Table 6.5).

Table 6.4: p -values for statistically significant pairwise comparisons of subjective ratings across test conditions.

Test condition	Test condition	p value
Visual only	Auditory only in noise at 5dB above bespoke SNR	.01
	Auditory only in noise at bespoke SNR	<.001
	Dual-tasking in noise at 5dB above bespoke SNR	<.001
	Dual-tasking in noise at bespoke SNR	<.001
Auditory only in quiet	Auditory only in noise at 5dB above bespoke SNR	.01
	Auditory only at bespoke SNR	<.001
	Dual-tasking in noise at 5dB above bespoke SNR	<.001
	Dual-tasking in noise at bespoke SNR	<.001
Auditory only in noise at 5dB above bespoke SNR	Dual-tasking in noise at bespoke SNR	.01
Dual-tasking in quiet	Auditory only in noise at bespoke SNR	.01
	Dual-tasking in noise at 5dB above SNR	.05
	Dual-tasking in noise at bespoke SNR	<.001

Table 6.5: p -values for statistically insignificant pairwise comparisons of subjective ratings across test conditions.

Test condition	Test condition	p value
Visual only	Auditory only in quiet	1.0
	Dual-tasking in quiet	.38
Dual-tasking in quiet	Auditory only in noise at 5dB above bespoke SNR	1.0
Auditory only in quiet	Dual-tasking in quiet	.30
Auditory only in noise at 5dB above bespoke SNR	Auditory only in noise at bespoke SNR	1.0
	Dual-tasking in noise at 5dB above bespoke SNR	1.0
Auditory only in noise at bespoke SNR	Dual-tasking in noise at bespoke SNR	1.0
	Dual-tasking in noise at 5dB above bespoke SNR	1.0
Dual-tasking in noise at bespoke SNR	Dual-tasking in noise at 5dB above bespoke SNR	1.0

It therefore appears that the subjective ratings are not in agreement with the LE scores obtained in quiet and in noise. Indeed, subjective ratings significantly increased when dual-tasking in noise (at either SNR) compared to when dual-tasking in quiet. This contrasts with the lack of significant differences in visual accuracy when dual-tasking in noise (compared to quiet). Furthermore, subjective ratings did not significantly increase when dual-tasking in quiet compared to the visual-only condition, which is contrary to the significant deterioration in visual accuracy when dual-tasking in quiet (relative to the visual-only condition).

If it is genuinely the case that there is no agreement between subjective ratings and visual performance, a potential explanation is that this particular type of subjective rating was not measuring the same aspect of LE being detected by the digit stream dual-task. This is a tenable notion since dissociations in data trends between subjective and behavioural indices of LE have been previously reported in LE research (McGarrigle et al., 2014). However, the CI users were exhibiting substantial variability in their ratings across the test conditions. This could then hinder detection of significant differences within the model.

Another factor to consider is that the participants may have performed incorrect resource allocation, hence no significant increase in LE in noise. Thus, there is a need to calculate primary task stability.

6.3.3.3. *Stability of primary task performance:* *Auditory accuracy*

An average was taken of the two runs for each test condition (see Appendix 6.3 for the auditory accuracy scores obtained in the two attempts for each condition). It emerged that the introduction of noise using the participants' bespoke SNRs had detrimentally affected the CI users' auditory accuracy (Figure 6.5 a). Furthermore, there were decreases in auditory accuracy when dual-tasking was required relative to that achieved within the auditory-only condition (Figure 6.5 b).

To ascertain the statistical significance of these trends, a 2x3 mixed effects logistic regression was applied to auditory accuracy percentage. Both predictors were categorical: task type (i.e. auditory-only or dual-task) and noise type (i.e. in quiet and in noise at the two different SNRs). Interaction was allowed in this model, and participants were also set as random effects.

6.3. Experiment 8

This model revealed an expected main effect of SNR: as SNR increased, auditory accuracy significantly deteriorated ($F(2,130)=388.6$, $p<.001$). The main effect of task, however, was not significant ($F(1,130)=2.3$, $p=.14$), indicating a lack of difference between auditory-only and dual-tasking conditions. There was also no significant interaction ($F(2,130)=0.3$, $p=.72$). Thus, there appears to be stability in primary task performance, i.e. resource allocation was likely to be appropriate (although the same earlier caveat still applies in relation to the interpretation of the statistically null results). This, in turn, suggests that the LE scores obtained in noise are likely to be valid. Therefore, the lack of change in visual accuracy in noise seems legitimate.

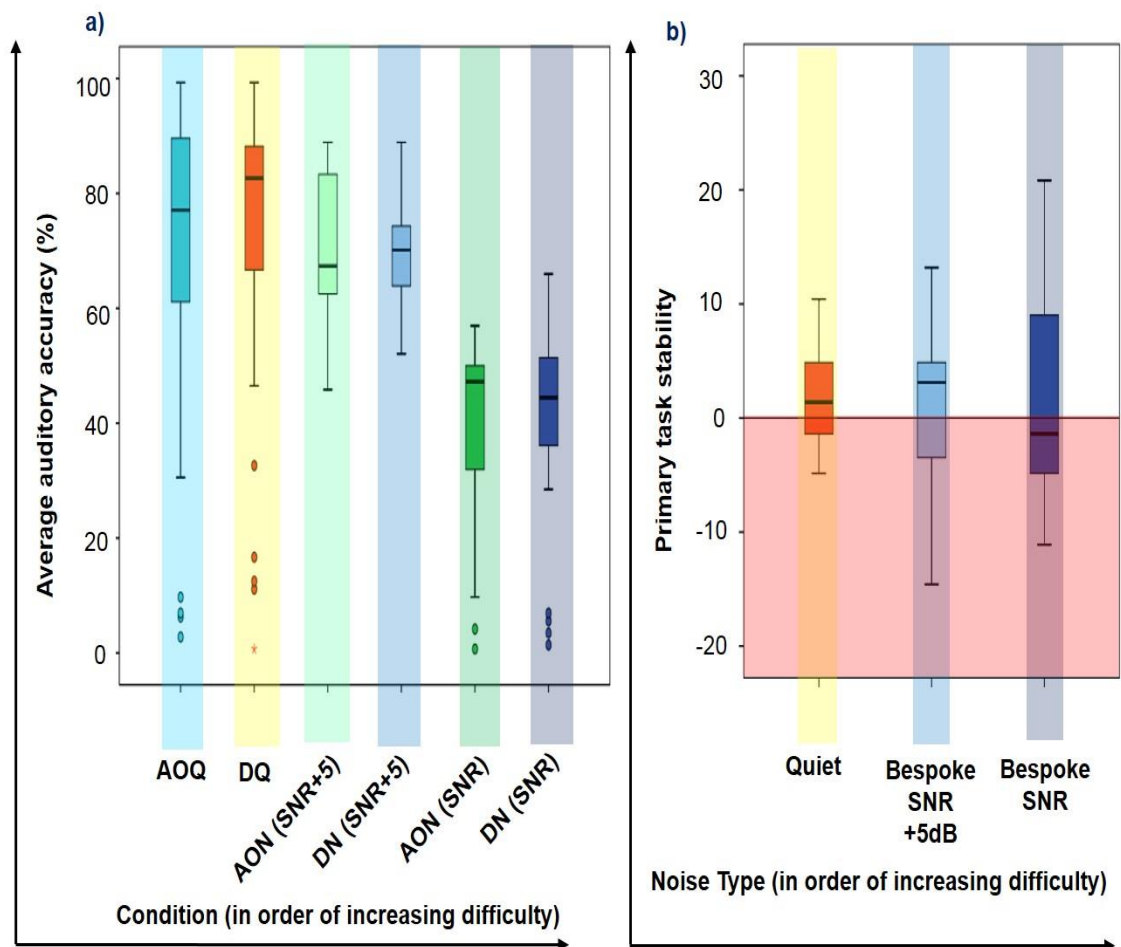


Figure 6.5: Boxplots of auditory accuracy (a) and of primary task stability (b).

AOQ = Auditory task only in quiet, DQ = Dual-tasking in quiet, AON (SNR) = Auditory task only in noise at bespoke SNR, DN (SNR) = Dual-tasking in noise at bespoke SNR, AON (SNR+5) = Auditory task only in noise at 5dB above bespoke SNR, DN (SNR+5) = Dual-tasking in noise at 5dB above bespoke SNR.

Red shading indicates deterioration in auditory accuracy when dual-tasking. Unit of primary task stability = Difference in percentage.

6.3.3.4. *Evaluation of viability of LE measurement*

Before any definitive conclusions can be drawn as to the reliability of LE measurement by digit stream dual-task (in both quiet and noise), the tracking data needs to be examined.

The tracking criterion had been tailored to each individual CI user, as part of the attempt to equalise the auditory performance produced by the two noise levels used in testing. There is a need to check that the adaptive tracking procedure did actually succeed in its objective in obtaining appropriate bespoke SNRs for the entire CI cohort.

It emerged that the majority of the CI users achieved reasonable tracks (as well as the expected sigmoidal psychometric functions: for exemplar, see the two attempts with affiliated psychometric functions and adaptive tracks produced by CI user 10 in Figure 6.6 b- e; for remainder of CI users' psychometric functions and tracks, see Appendices 6.4-6.16).

Furthermore, when test-retest reliability of the noise tracking (across the two tracking attempts) was assessed using the intraclass correlation coefficient (ICC), a relatively high level of agreement was found (ICC=0.76). This thus suggested a certain amount of consistency in auditory performance during the adaptive tracking.

In addition to creating bespoke tracking criteria, a minimum of 35% was set to the tracking level applied. This was due to the concern that tracking any lower level would essentially be meaningless (because no granularity would be present at this floor level of performance). However, this minimum tracking level had the consequence that, for seven participants, their tracking criterion was not actually set at 50% of the achieved performance in quiet, because their accuracy (in quiet) was below 70% correct (see data points lying outside shaded red oval of Figure 6.6 a).

With these seven participants, tracking was still pursued but with the tracking criterion being set at the minimum (i.e. 35% correct). This was in case the obtained performances in quiet were not actually representative of the participant's true speech intelligibility performance (i.e. it was actually possible to track these participants).

For one of the participants, this was indeed the case within their second attempt of tracking (with the criterion set at 35%) yielding the SRT of 10.9dB SNR (i.e. CI user 20: see Figure 6.6 f- i).

However, for the remaining six participants, tracking failed (for exemplar of failed attempt, see Figure 6.8 j and k as produced by CI user 14).

Despite these failures in tracking, testing in noise was still pursued, but with the SNR set at the maximum (i.e. 20dB). The original intention of this was to confirm whether the digit stream dual-task was still capable of detecting changes in LE in these CI users, despite such poor auditory performances.

It also needs to be noted that, for those who failed tracking, there was no longer application of test conditions where the noise level was set at 5dB above the bespoke SNR (due to the maximum SNR already being used).

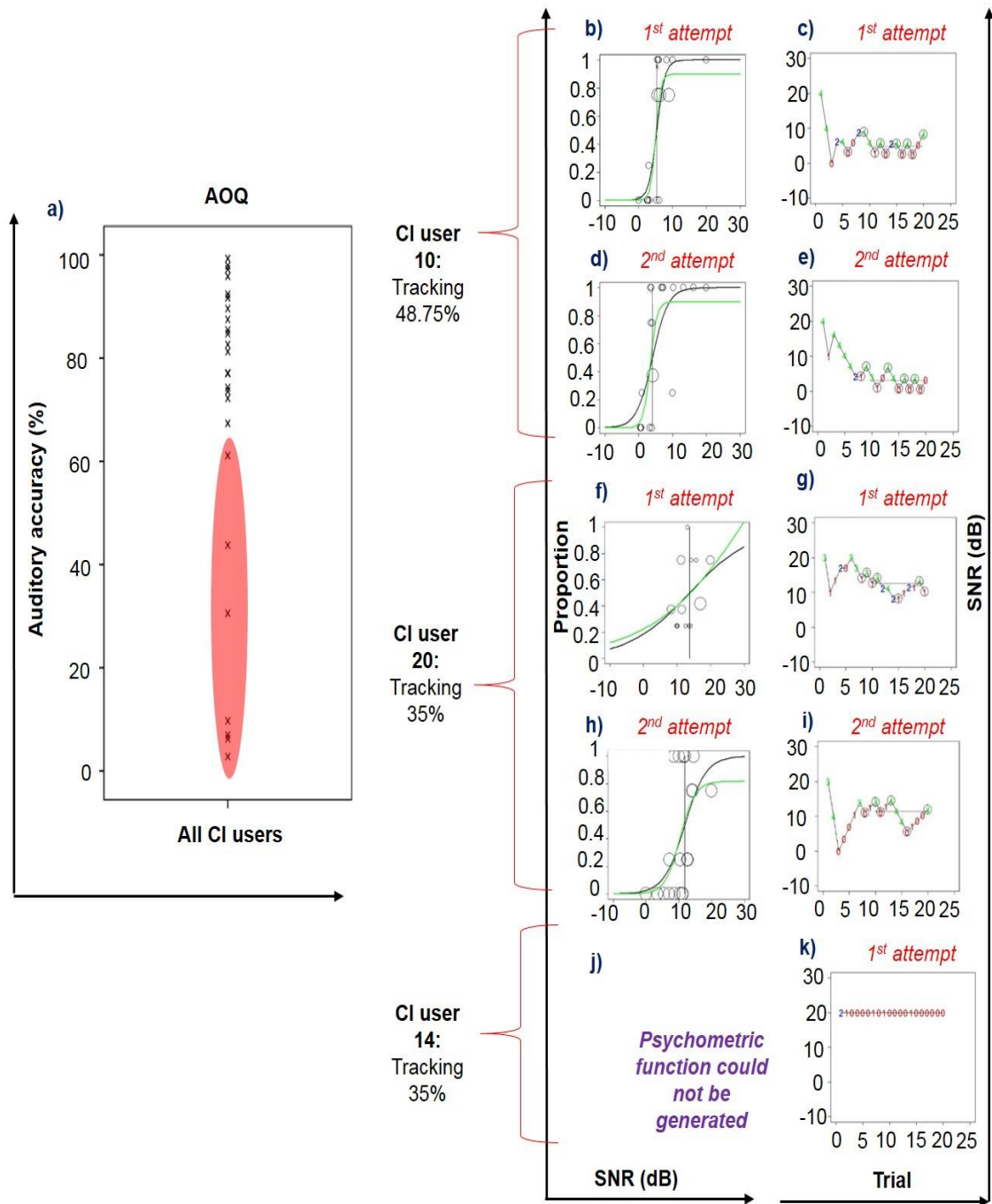


Figure 6.6: Dot plot of auditory accuracy in quiet (a) and exemplars of psychometric functions and affiliated adaptive tracks (b – k).

The psychometric functions (b, d, f, h, j) depict the proportion of average number of key words successfully heard at a given SNR. The adaptive tracks (c, e, g, i, k) depict how SNR was changed across the trials according to the adaptive algorithm.

Black line on psychometric function = standard logistic regression. Green line on psychometric function = logistic regression with additional parameter in the fit, i.e. an upper asymptote, otherwise known as lapse rate (to enable psychometric function to plateau at $y < 1$). Superimposed circles on psychometric function = Participant's response data, with size of circle controlled by the number of trials run at that particular SNR. Superimposed circles on adaptive track = Reversals. Superimposed numbers on adaptive track = Total number of key words correct, with colour indicating whether the track remained at the same SNR (blue), increased (red), or decreased (green) on the next trial.

Thus far, all data plots displayed in this chapter (i.e. visual, auditory and subjective ratings data) have included these CI users who could not be tracked. There is thus the possibility that by the inclusion of these participants may have skewed the group-level data. This is because it could not be guaranteed that using the SNR of 20dB was actually appropriate for all these CI users. Furthermore, these six CI users' performance in quiet would be relatively uncharacteristic of the overall participant group (being considerably poorer).

Therefore, the subjective ratings, visual accuracy scores and auditory performances will be reconsidered, with the datasets being parsed into "successes" (i.e. the CI users who could be tracked) and "failures" (i.e. the particularly poor performers who failed tracking).

Upon the parsing of the subjective ratings, it was revealed that there was extensive overlap in the scores, particularly within the visual only condition as well as the dual-tasking in noise with SNR set at the SRT (Figure 6.7). However, there did appear to be some negative skewing of the group data by the "failures" within both the auditory-only and dual-tasking conditions in quiet. In addition, the inclusion of the "failures" seemed to increase the spread of ratings given (i.e. towards both extremes of the scale) for the auditory-only condition where noise level was set at the bespoke SNR. However, this increase in spread had no influence on the group median. Therefore, in general, the influence of the "failures" on the nature of subjective ratings was relatively minimal.

The impact caused by the inclusion of the "failures" also seemed to be minimal for the visual accuracy scores (Figure 6.8 a- c), and also for the LE scores (Figure 6.8 d- f). The only observation of note was that, within the LE scores in the quiet listening condition, the "failures" seemed to cause both the range and interquartile range to increase.

This absence of effect by the "failures" (in both the subjective ratings and visual performances) was in spite of the obvious inferiority exhibited by the "failures" within their auditory accuracy across all test conditions (Figure 6.9 a- c). Indeed, when the contribution of the "failures" within the primary task stability calculation was explored, there appeared to be no influence from the "failures" (Figure 6.9 d- f).

Thus, despite the issues of tracking potentially jeopardising accurate LE measurement for the "failures", this does not appear to have had a detrimental impact at the group-level.

6.3. Experiment 8

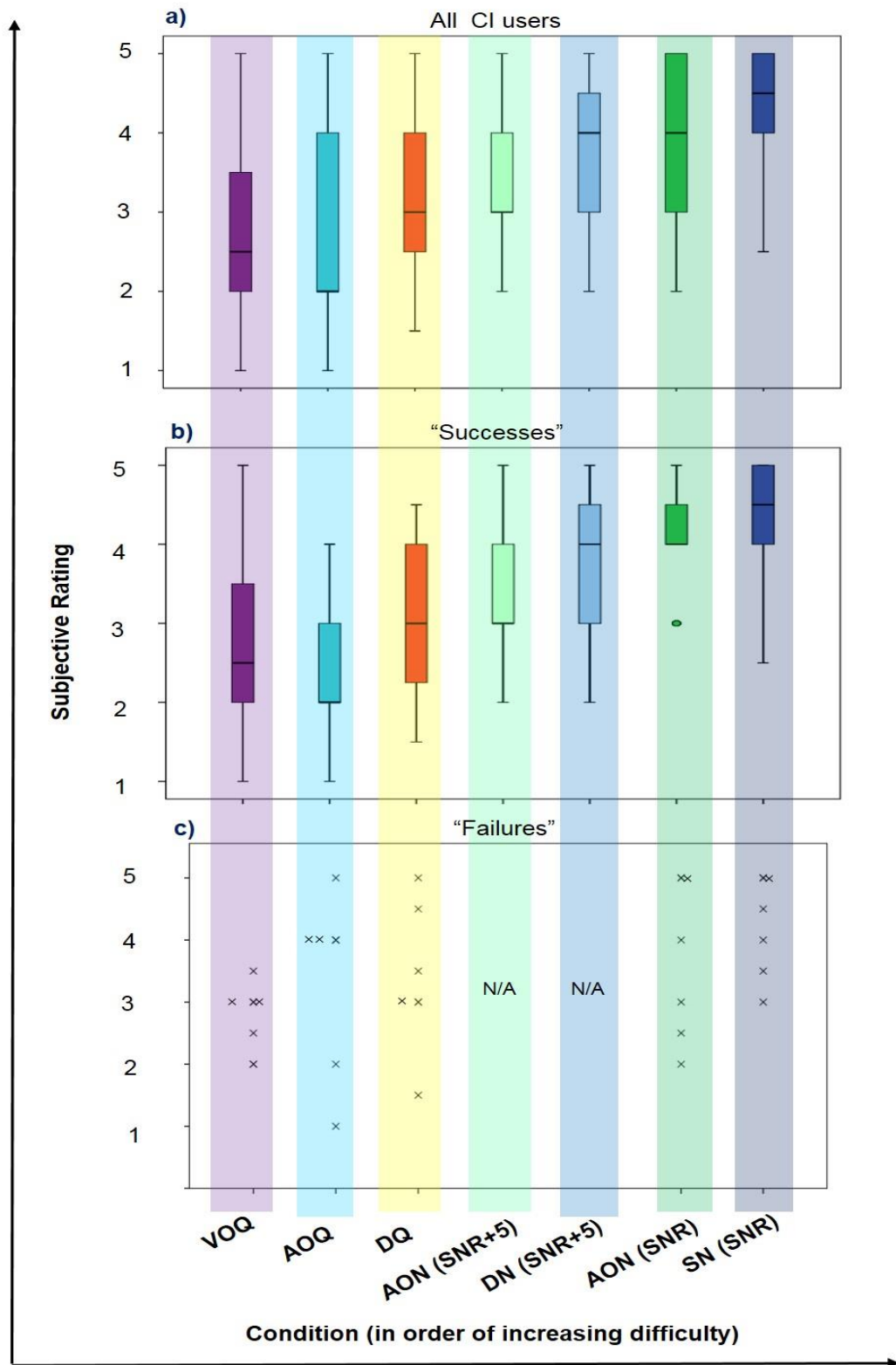


Figure 6.7: Boxplots of subjective ratings for the entire CI cohort **(a)** for “successes” who completed tracking **(b)**, and dot plot of subjective ratings for the “failures” **(c)**. VOQ = Visual task only in quiet, AOQ = Auditory task only in quiet, DQ = Dual-tasking in quiet, AON (SNR) = Auditory task only in noise at bespoke SNR, DN (SNR) = Dual-tasking in noise at bespoke SNR, AON (SNR+5) = Auditory task only in noise at 5dB above bespoke SNR, DN (SNR+5) = Dual-tasking in noise at 5dB above bespoke SNR. [N.B. The “failures” did not perform the conditions of AON (SNR+5) and DN (SNR+5).]

Subjective ratings: 1= Not at all hard work, 2 = Quite hard work, 3 = Medium hard work, 4 = Very hard work, 5 = Extremely hard work.

6.3. Experiment 8

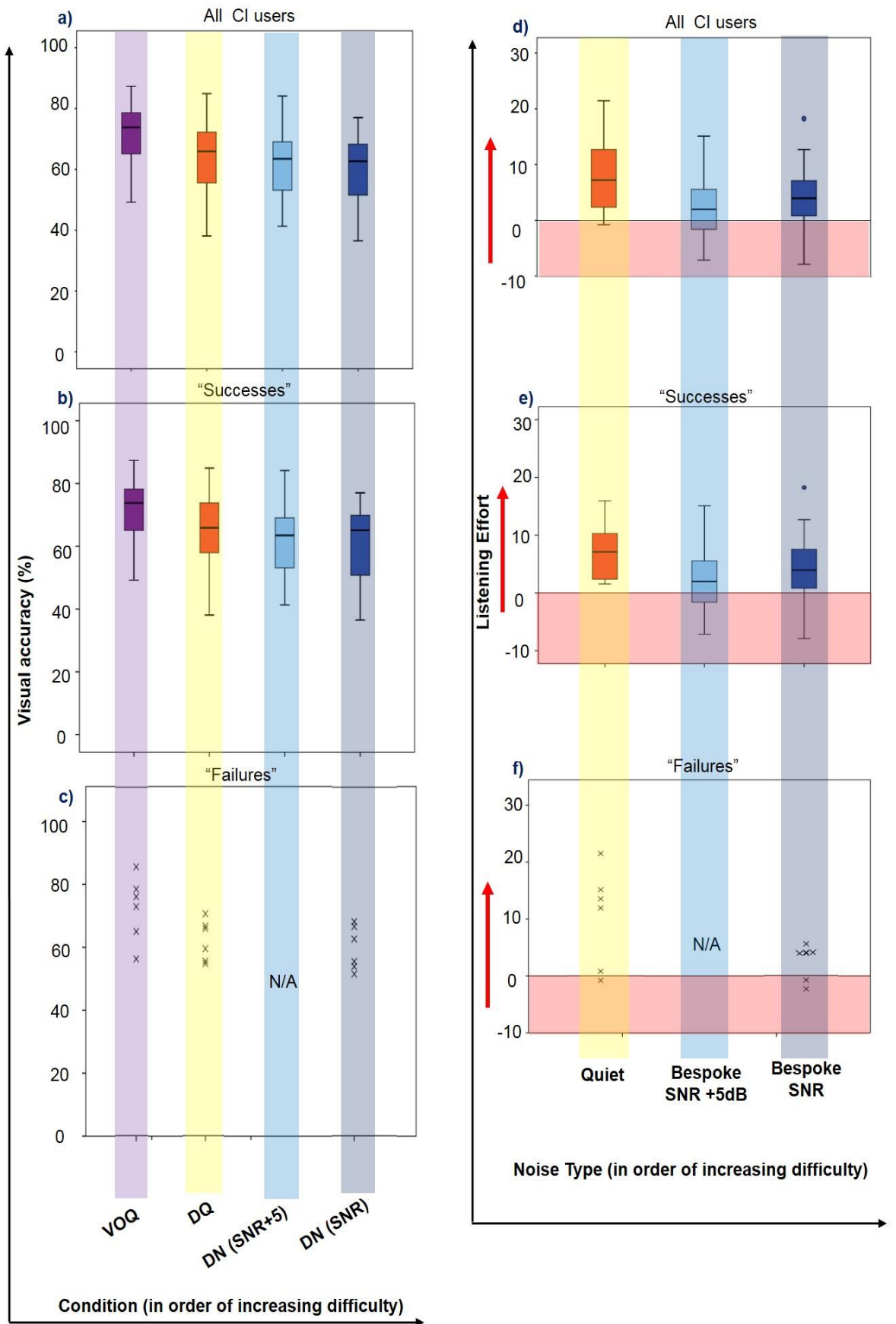


Figure 6.8: As for Figure 6.7, but displaying visual accuracy (a, b, c) and LE (d, e, f). Red arrow indicates direction in which LE increases. Red shading indicates decrease in LE. Unit of LE = Difference in percentage.

6.3. Experiment 8

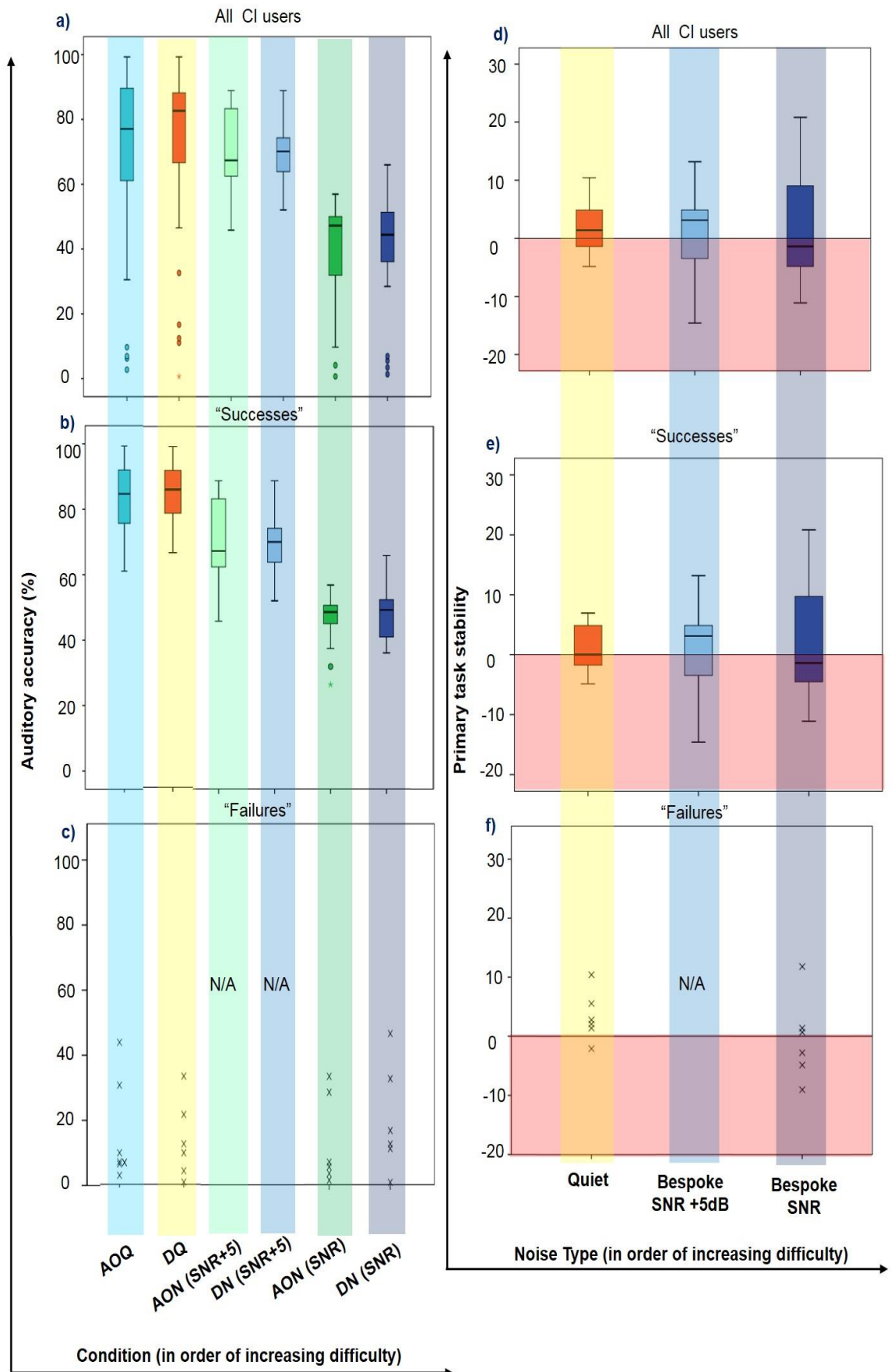


Figure 6.9: As for Figure 6.7, but displaying auditory accuracy (**a, b, c**) and primary task stability (**d, e, f**). Red shading indicates deterioration in auditory accuracy when dual-tasking. Unit of primary task stability = Difference in percentage.

It is also necessary to establish whether the tracking of the “successes” was actually appropriate. Because none of the tracks (or psychometric functions) attained by the “successes” were obviously aberrant, the best way to review this data is to consider the obtained SNRs within the context of the auditory performances obtained during testing.

Accordingly, the intended performance level (i.e. as defined by the tracking criterion applied) was compared to auditory accuracy achieved when the participant performed the auditory-only condition with noise set at their bespoke SNR (Figure 6.10).

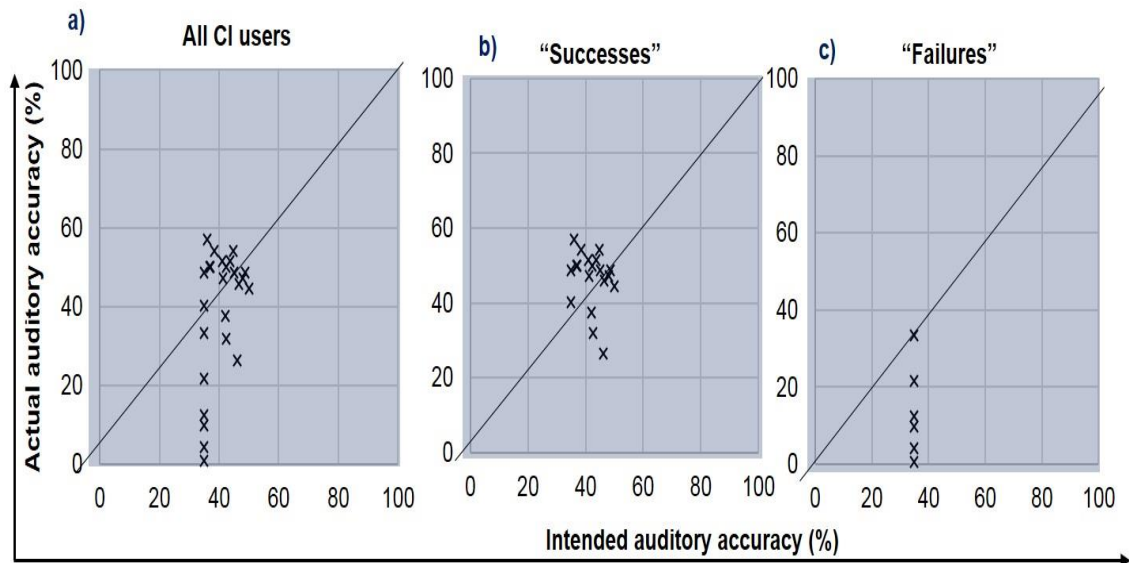


Figure 6.10: Scatterplots of actual versus intended auditory performance (of the auditory-only condition with noise set at the participants’ bespoke SNR) by the entire CI cohort **(a)**. The same dataset is broken down into the “successes” who completed tracking, and the “failures” **(b, c)**.

It was discovered that very few of the participants actually achieved the performance level intended by the tracking. Indeed, when only the “successes” were examined, there was no significant correlation: Pearson’s $r=0.441$, $p=.027$ (significance level: $p=.01$). Thus, there was the possibility that some (or, indeed, all) of the bespoke SNRs produced by the adaptive tracking for the “successes” do not represent the CI users’ true ability for speech perception in noise. There is now a concern that the acoustic challenge was not appropriate for the “successes”, let alone the “failures”. This may then account for the lack of significant LE detected in noise.

However, there is another aspect of the data that also needs to be interrogated: the nature of visual response. This is in order to determine whether over-reporting continues to be an issue.

As before, the number of trials for each type of visual response were tallied up across test conditions and target presentation frequency (for each test run). It emerged that, as the number of targets presented in the digit stream increased, the participants' visual response tended to drift towards that of reduced target detection (Figure 6.11). This was as expected. Unfortunately, it was also revealed that over-reporting was still occurring.

Therefore, the instances of over-reporting were isolated from the dataset for each test run, across all test condition and target presentation frequency (Figure 6.12). It became apparent that all CI users, regardless of they were a "success" or a "failure", were susceptible. The maximum number of trials with over-reporting was 3.

Not only this, there appeared to be the tendency that, as the test condition increased in difficulty, the incidence of over-reporting decreased (Figure 6.13a). In addition, as the number of targets presented in the digit stream increased, over-reporting incidence also decreased (Figure 6.13b).

To ascertain the statistical significance of the effect of test condition and target presentation frequency on over-reporting, the over-reporting incidence was averaged and a mixed effects logistic regression model was generated. There were two predictors: test condition, which was treated as a categorical variable (i.e. visual only, dual-tasking in quiet at the two levels of spectral resolution and dual-tasking in noise), and number of targets (presented in the digit stream), which was treated as continuous. Interaction was allowed, and participants were set as random effects.

This model revealed a significant main effect of test condition ($F(3,364)=2.04$, $p=.11$), as well as a significant main effect of target presentation frequency ($F(1,364)=64.4$, $p<.001$). Thus, over-reporting significantly decreased when SNR deteriorated, and also when the number of targets increased. However, there was no significant interaction ($F(3,364)=2.5$, $p=.06$), indicating that the decrease in over-reporting with target number was similar for each SNR.

Since over-reporting was still occurring, and in a systematic manner, this adds to the concern that the LE scores may be unreliable. This is because instances of over-reporting may compromise visual accuracy measurement, which is the basis of the LE calculation.

6.3. Experiment 8

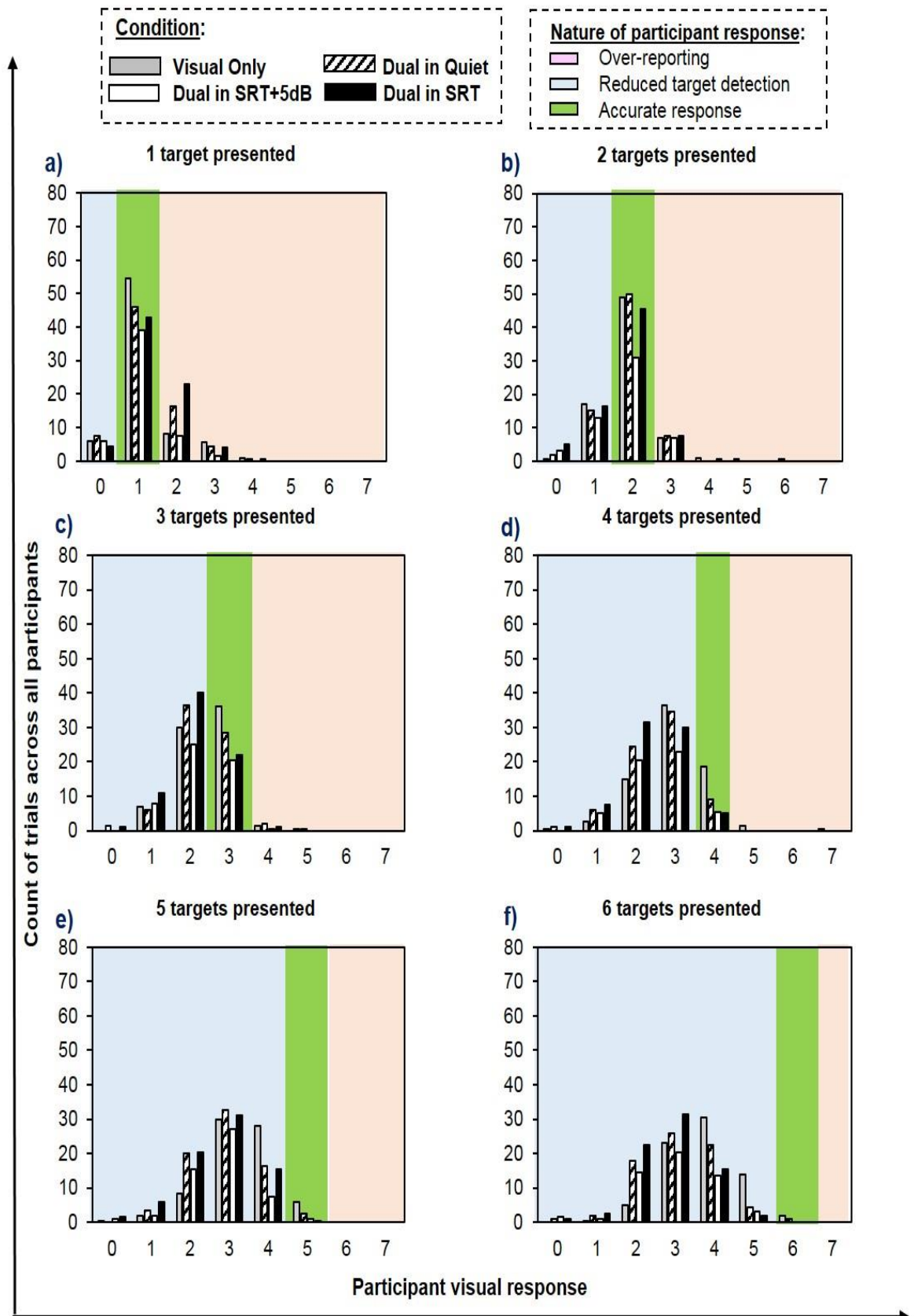


Figure 6.11: Histograms depicting the nature of participants' visual responses across trials (in each test condition) for each target presentation frequency.

6.3. Experiment 8

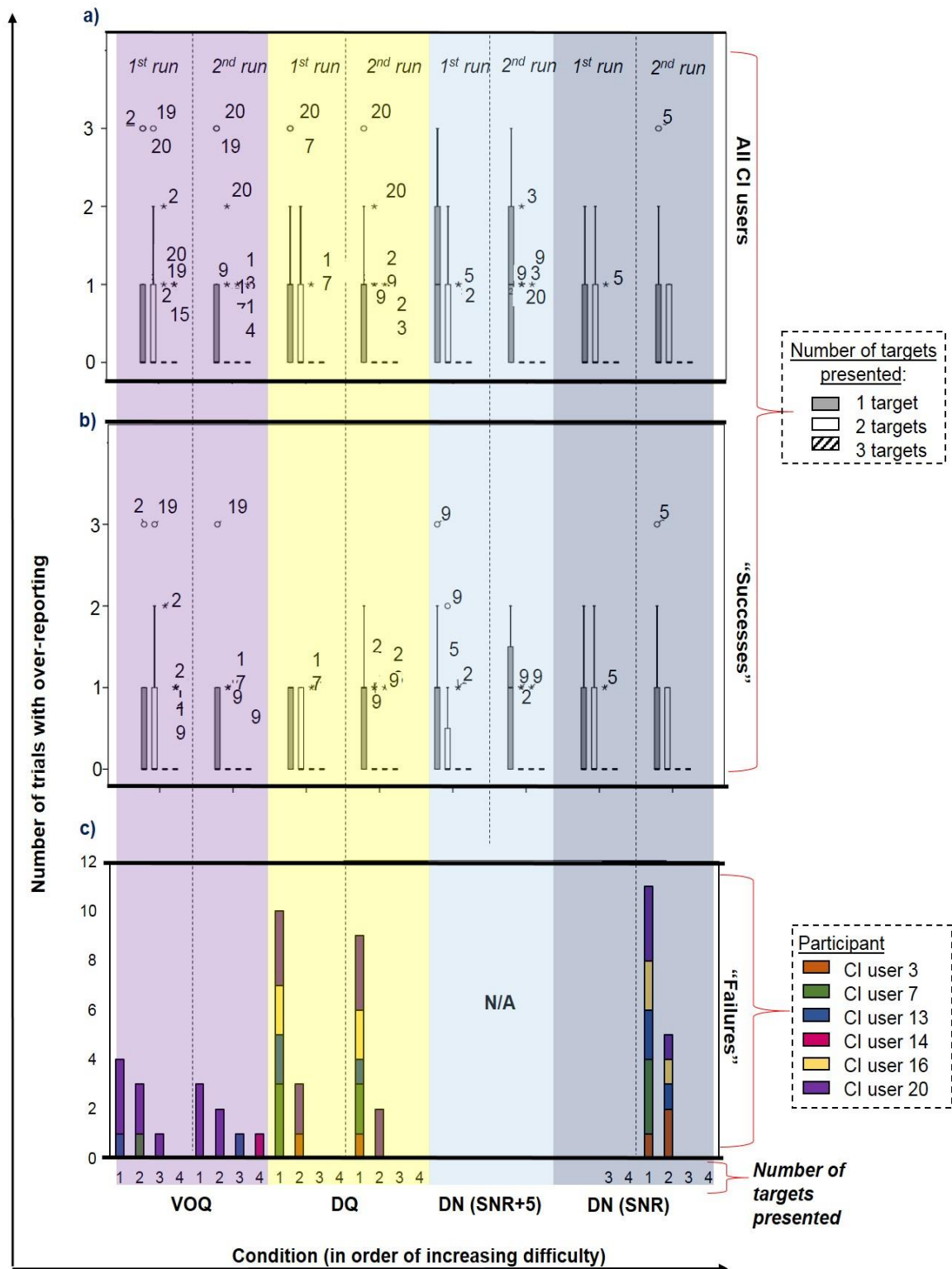


Figure 6.12: Boxplots and stacked histograms showing incidence of over-reporting according to the number of targets actually presented in the digit stream (in terms of number of trials within given test conditions). **(a)** All CI users. **(b)** "Successes" only (CI users who successfully completed tracking). **(c)** "Failures" only (CI users who failed to track). VOQ = Visual task only in quiet, DQ = Dual-tasking in quiet, DN (SNR) = Dual-tasking in noise at the participant's bespoke SNR, DN (SNR+5) = Dual-tasking in noise at 5dB above bespoke SNR. Asterisks (*) indicate that the occurrence of that particular value was an outlier. Superimposed numbers relate to the numerical identity of NH control(s) responsible for the given outlier value.

6.3. Experiment 8

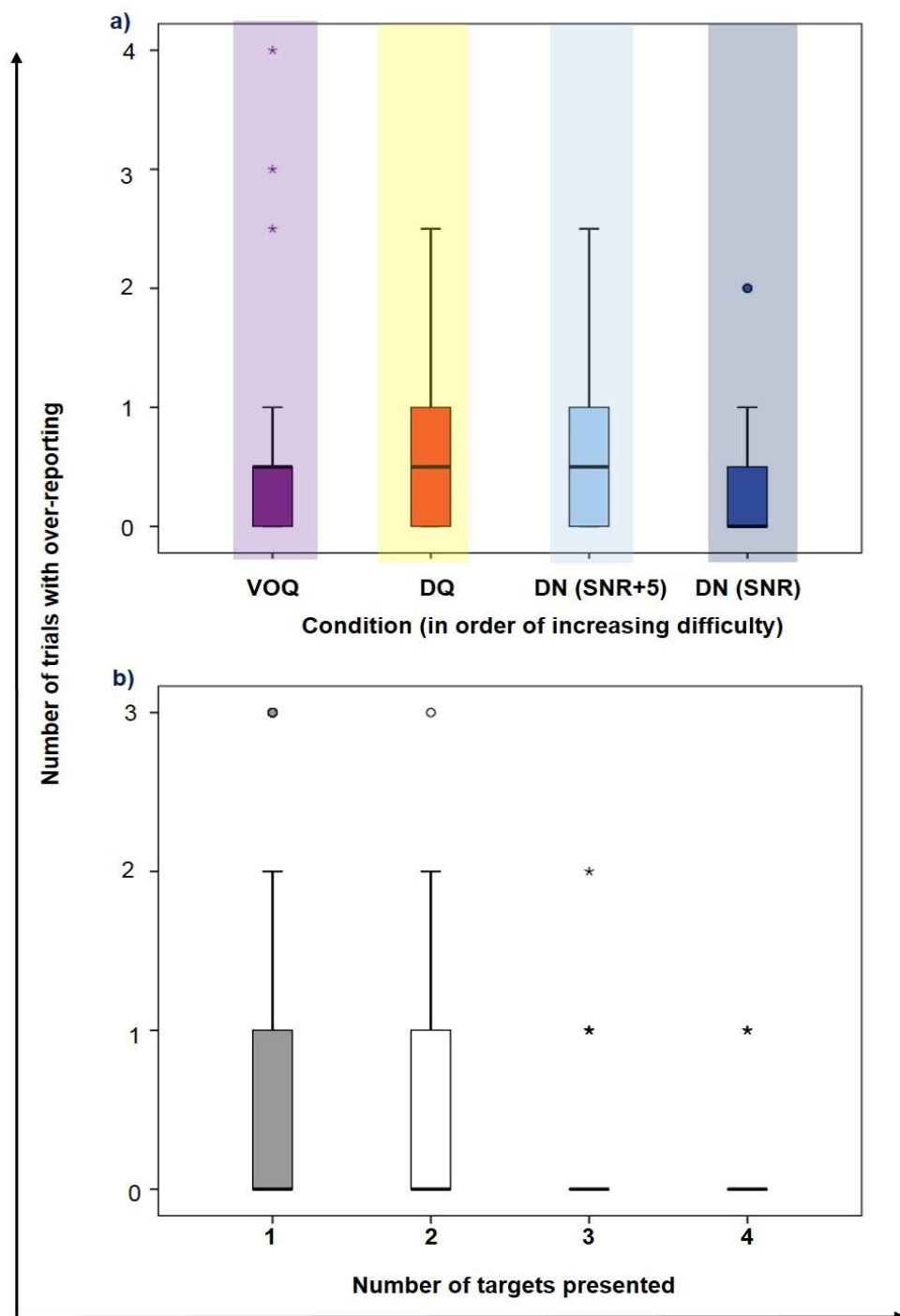


Figure 6.13: Boxplots for each test condition (summed across all conditions of target presentation frequency: **b**), and for each target presentation frequency (averaged across all test conditions: **c**). Total number of trials for each test condition is 18.

VOQ = Visual task only in quiet, DQ = Dual-tasking in quiet, DN (SNR) = Dual-tasking in noise at the participant's bespoke SNR, DN (SNR+5) = Dual-tasking in noise at 5dB above bespoke SNR. Asterisks (*) indicate that the occurrence of a particular value was an outlier. Circles indicate extreme values.

6.4. Conclusions of Experiment 8

There appears to be promise that primary task stability was achieved, suggesting appropriate resource allocation had occurred during testing. The digit stream dual-task detected statistically significant increase in LE when the CI users were listening in quiet. In noise, the LE scores approached significance ($p=.056$). Borderline significance was achieved in LE at the CI users' bespoke SNR, but no significance in LE was found at the easier SNR ($p=.4$). However, this does not necessarily mean that listening in noise was effortless, but simply that no statistically significant difference could be detected in visual accuracy between dual-tasking in quiet and dual-tasking in noise.

Indeed, the subjective ratings did not corroborate the visual accuracy scores, with significant increases in rating of perceived effort when comparing dual-tasking in noise to dual-tasking in quiet. Interrogation of the tracking data also identified a lack of correlation between bespoke SNR and performance in auditory accuracy. This suggests that acoustic challenge presented by noise may not have been appropriate for all CI users. It needs to be noted that the inclusion of the "failures" (i.e. CI users who could not be tracked at all) did not seem to influence group effects.

In addition, over-reporting persisted with significant main effects, which was not desirable. This is because over-reporting may affect the reliability and validity of the visual accuracy measurement, which is an integral part of the calculation of LE. This, in turn, potentially compromises the feasibility of the digit stream dual-task as a framework for the measurement of LE.

PART FIVE:

DISCUSSION

CHAPTER SEVEN:

General discussion and future directions

7.1. Introduction

The overall aim of this thesis is to evaluate the feasibility of using the behavioural framework of the dual-task paradigm as the basis for a clinical test of LE in CI users. Accordingly, the experiments in this thesis have attempted to interrogate the ability of the dual-task paradigm to detect and quantify LE. Over the course of the experiments, specific elements of the dual-task design have been systematically modified in response to the findings of each individual experiment. These refinements were intended to optimise the dual-task's sensitivity to the LE construct, in order to potentially achieve the goal of feasibility.

7.2. Summary of all experiments

There were nine experiments in total. A recall dual-task was initially tested within the Pilot Study, which was then adapted into a digit stream dual-task for Experiments 1-8. This was because of the need to more specifically control the task difficulty level.

The digit stream dual-task itself was then carefully fine-tuned during its trialling with normal hearing participants within two different listening conditions: normal and CI simulations. The modifications undertaken were essentially systematic troubleshooting of methodological issues as they arose. Such issues included the potential interference of retinal image persistence (Experiment 1); individual differences in speech-in-noise processing causing inequality in testing conditions between the participants (Experiment 3); as well as continued perceptual learning of novel auditory stimuli skewing auditory performance (Experiment 5). Accordingly, the backward masking technique (Experiment 2 onwards), adaptive tracking (Experiment 4 onwards) and an extended training regime for the novel auditory stimuli (Experiments 6 and 7) were all implemented respectively in the ongoing attempt to refine the dual-task.

Each of these alterations in dual-task design was ultimately guided by three cardinal tenets of cognition: 1) Cognitive resources are required for task execution; 2) Cognitive capacity is finite; and 3) These limited cognitive resources have to be shared between multiple tasks (Pichora-Fuller et al., 2016; Edwards, 2016).

7.3. Summary of key findings

The dual-task paradigm exploits this finite cognitive capacity by demanding simultaneous execution of two tasks, in order to use resource allocation to infer the listening effort (LE) expended. However, the dual-task paradigm can only be successful in indexing LE if resource allocation is appropriate: i.e. performance of only the secondary task deteriorates, due to the prioritisation of cognitive resource for the primary task (at the cost of the secondary task). Therefore, in all experiments, the analysis always involved the determination of whether this desired pattern in primary task stability and deterioration in secondary task performance had actually occurred.

In the final experiment (i.e. Experiment 8), the viability of the digit stream dual-task as a form of LE measurement was tested with the target clinical population, i.e. the CI users.

7.3. Summary of key findings

The key experiments, which implemented the modifications of the earlier studies were Experiments 7 and 8.

In Experiment 7, the channel tracking appeared to be successful in achieving a bespoke spectral resolution for CI simulations in NH participants. Under these conditions, the LE in quiet increased when the spectral resolution was set at the bespoke channel number. When the noise level was fixed at 40-65% correct, a further significant increase in LE was also detected. However, these changes in LE scores were only applicable when the spectral resolution of auditory stimuli was between 4-8 channels (and not higher). It has been previously noted that 8 channels is representative of the typical spectral resolution achieved in a high-performing CI user (Friesen et al., 2001), so 8 channels or fewer is potentially relevant to the acoustic experience of a range of less able CI users.

These increases in LE were reflected in the subjective ratings of Experiment 7, which followed the same trends. These results thus suggest that the LE imposed by the degraded auditory signal (delivered by the CI simulations) was sizeable, and it increased still further when noise was introduced.

These results in quiet were supported by the findings of Experiment 8 involving CI users. Once again, LE was found to be significantly increased in quiet, as inferred from visual accuracy decreasing when dual-tasking in quiet compared to the visual-only condition. However, the increase in LE in noise just failed to reach significance. Indeed, the increase in LE at the bespoke SNR was found to be of borderline significance, and no significant increase was identified at the easier SNR.

Interestingly, these borderline results in noise were not corroborated by the subjective ratings, which displayed a significant increase when dual-tasking in noise compared to dual-tasking in quiet. This suggests that the CI users had perceived listening in noise as substantially harder than listening in quiet. To complicate matters, further analysis detected a lack of correlation between auditory performance and the imposed tracking, indicating that the tracking procedures may not have necessarily equalised the level of difficulty in the noise condition for all CI users.

Testing in noise, therefore, proved to be particularly problematic for the CI users. However, the most striking result was the increase in LE in quiet conditions, which was also reflected in the NH controls listening to CI simulations (i.e. Experiment 7). This is an unexpected outcome for the CI users. It was hypothesised that LE would increase in noise, but there had been no expectation that significantly increased levels of LE would already be present in the optimal listening condition of quiet.

7.4. Application of key results to cognitive capacity

The core tenets of the dual-task paradigm appear to be operating correctly in Experiments 7 and 8 (i.e. CI simulations and CI user studies). Primary task stability was maintained (i.e. there was no significant change in auditory accuracy in the single-task and dual-task conditions). Also, deterioration only occurred in the secondary visual task. This suggests that the primary task was being appropriately prioritised, leaving fewer cognitive resources to address the visual task, leading to poorer visual accuracy. All these data trends indicate that the design of the digit stream dual-task was fit for purpose.

Using the idea of finite cognitive resources as a model, it was initially hypothesised that performing the auditory task in quiet would be of a low cognitive load, i.e. relatively few cognitive resources would be needed for successful auditory performance. Thus, when the participant is dual-tasking, the residual cognitive capacity left for the visual task should be sufficient for successful visual performance too (i.e. visual accuracy achieved when dual-tasking is no different to that achieved when performing just the visual task: Figure 7.1).

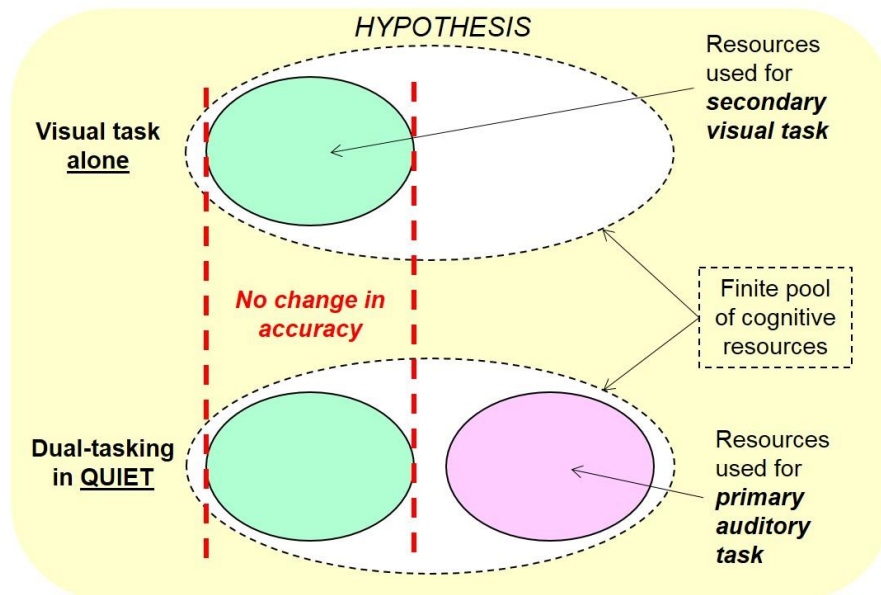


Figure 7.1: Venn diagrams showing the theoretical allocation of cognitive resources to primary and secondary tasks when performing just the visual task compared to dual-tasking in quiet.

Thus, the fact that visual accuracy significantly deteriorated in quiet ($p < .001$) when CI simulations were applied (compared to performing the visual task alone) suggests that considerably more cognitive resources had to be allocated to the auditory task than originally anticipated (Figure 7.2).

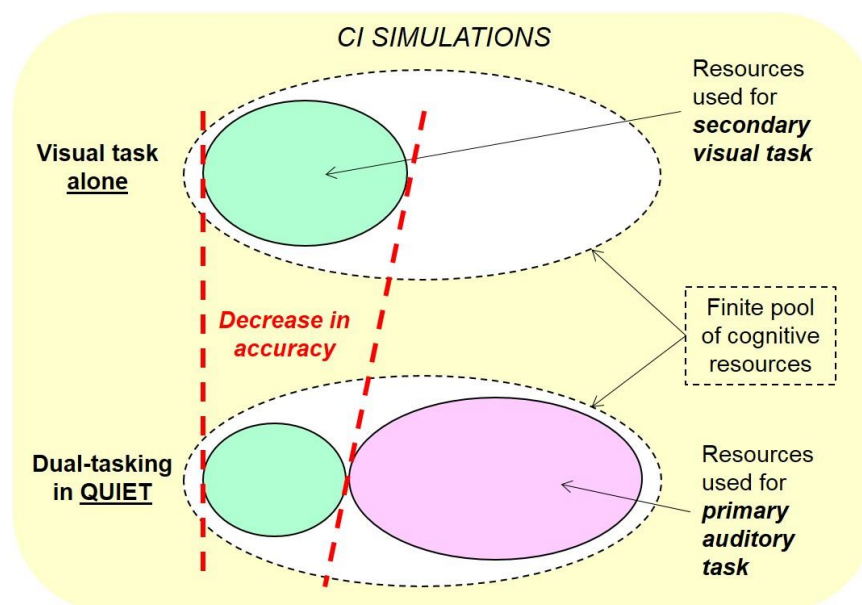


Figure 7.2: As for Figure 7.1, but representing the CI simulations case.

However, the exact quantity of additional cognitive resources needed for listening in quiet cannot be specifically measured, because the current LE calculation cannot be translated into a graduated scale.

When noise is introduced, even more cognitive resources are re-allocated to the auditory task (although this movement of resources is relatively smaller). This causes a further decrease in residual capacity left for the visual task (Figure 7.3). Accordingly, a significant deterioration in visual accuracy is detected ($p=.021$).

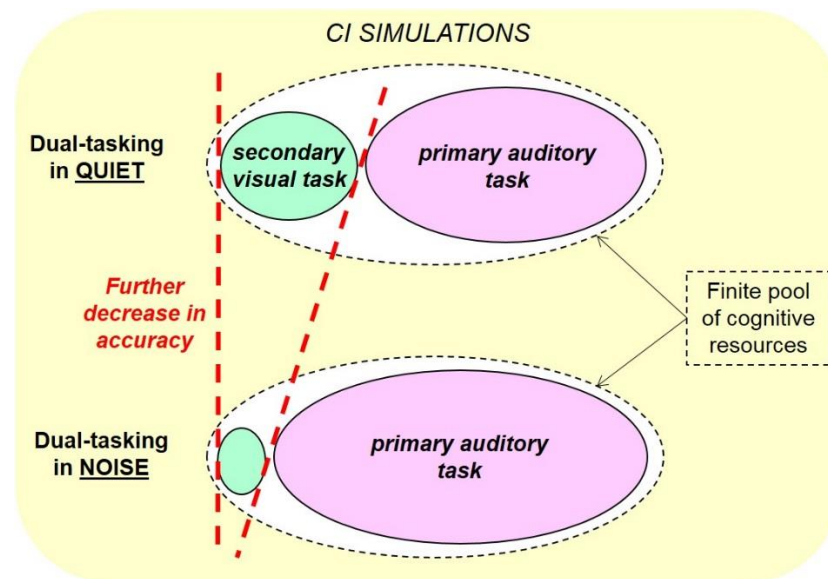


Figure 7.3. As for Figure 7.2, but for test conditions of dual-tasking in quiet and noise.

In the case of the CI users, it was notable that deterioration in visual accuracy was barely significant when dual-tasking in noise ($p=.056$). This was in spite of significant increases of LE in quiet. An explanation could be that an even greater amount of cognitive resources was required by the CI users to perform the auditory task in quiet (compared to that necessitated by the CI simulations for the NH controls: Figure 7.4).

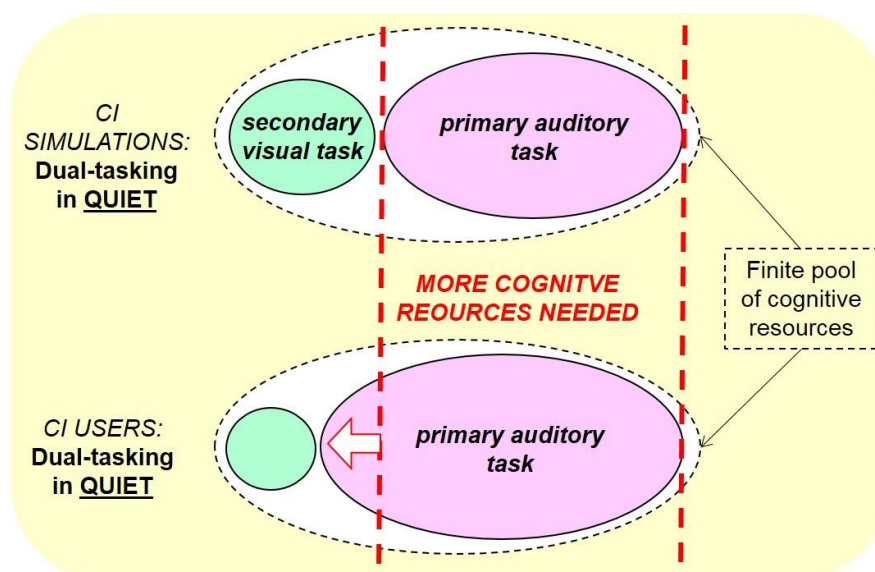


Figure 7.4. Venn diagrams comparing the theoretical allocation of cognitive resources when dual-tasking in quiet for CI simulations and actual CI users.

This increased cognitive load could potentially be incurred by the additional disadvantages imposed by the deafened brain (Kral et al., 2016; O'Donoghue et al., 2016). Accordingly, when noise was introduced, the margin for re-allocation of cognitive resources is even smaller than that available within the case of CI simulations (Figure 7.5). This then means that the detected changes in cognitive resources become less likely to reach statistical significance. Indeed, it was only the hardest SNR that reached borderline significance ($p=.05$), unlike the easier SNR ($p=.4$).

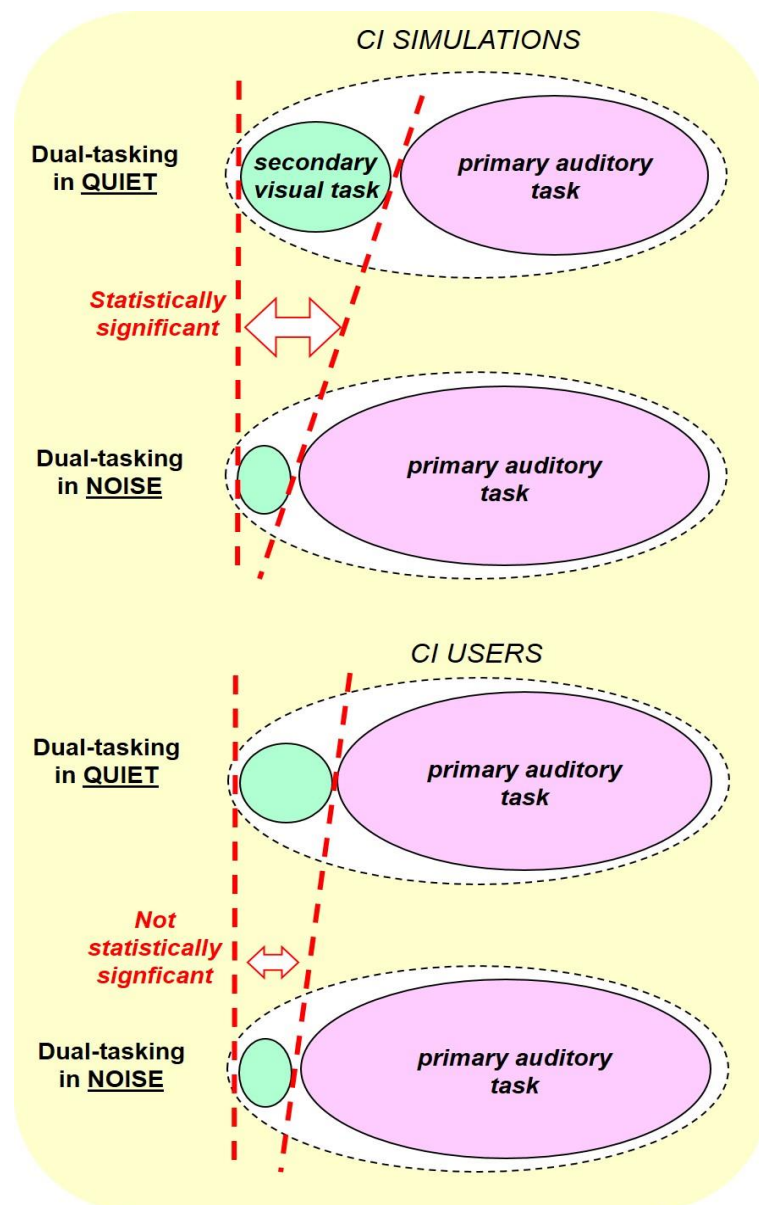


Figure 7.5. Venn diagrams comparing the theoretical allocation of cognitive resources when dual-tasking in quiet and in noise for CI simulations and actual CI users.

Ultimately, the CI users did not manage well in noise, potentially because they had already allocated the majority of their cognitive resources to simply processing the CI input, even before any challenging listening conditions were introduced. This might then explain why the tracking procedures in noise proved so problematic for the CI cohort. With such a small cognitive capacity left to cope with the noise, it would be very difficult to introduce sufficient gradation (in SNR), in order to successfully generate “hard” and “easy” conditions of noise (at least with the current adaptive tracking algorithm).

It is even possible that such gradation in tracking might not be feasible. Indeed, it could be the case that, for a CI user, noise is just noise. In other words, noise presents an immediate universal challenge, irrespective of its severity.

7.5. Relationship between key results and existing literature

Evidence from a wide range of studies assessing the impact of noise-excited vocoding via behavioural or physiological measures (including pupil dilation and neuroimaging techniques e.g. MRI and fNIRS) have all suggested that listening to spectrally degraded speech is effortful, compared to when processing clear speech (Pals et al., 2013; Winn et al., 2015; Wild et al., 2012; Wijayasiri et al., 2017).

However, there is currently a lack of empirical data concerning exactly what happens in the deafened brain of the skilled CI user, particularly when listening in optimal listening conditions (i.e. quiet). This is a striking omission, particularly in light of the evidence related to the physiological changes that occur in the deafened brain, and the impact of deafness as a connectome disease (Kral et al., 2016; O'Donoghue et al., 2016).

The key issue is that noise-excited vocoding studies tend to involve temporary exposure of a hearing brain to degraded sound, as well as only brief periods of training (in terms of hours, rather than days). For the CI user, however, exposure to the impoverished auditory input (that the CI device provides) begins from the moment the CI processor is switched on, and continues on for years and even decades (in the case of CI users implanted as children maturing into adulthood). Accordingly, it is easy (and not unreasonable) to assume that some form of acclimatisation and/or adaptation will eventually occur. This, in turn, can lead to the expectation that the CI user's auditory processing (of the CI input) will become increasingly more efficient over time.

However, the data from this present thesis suggest that even successful and skilled CI users (averaging 13 years of experience with the CI device) exhibit increased LE levels even in optimal listening conditions of quiet.

This effortful listening in quiet may arise from compensatory neurocognitive processes, such as perceptual filling-in or phonemic restoration (Başkent, 2014; Mattys et al., 2012). It is possible that these compensatory mechanisms are simply inevitable processes which are generated in response to the impoverished nature of the CI input.

If this is the case, the requirement to fill in the gaps created by the missing auditory information (in the CI input) may create a lifelong drain on cognitive resources even in the easiest of listening conditions possible, and irrespective of the length of wear for the CI users.

7.6. Potential implications of key results

If additional cognitive load (potentially required by compensatory mechanisms) occurs as soon as the CI device is switched on, this has implications regarding “wear time” for the CI user. This wear time is currently clinically monitored via data logging. Most CI processors have data logging as a feature, meaning that clinical audiologists in the auditory implant programmes are able to monitor the CI system use (Cochlear Americas, 2013; Cochlear, 2015). This involves an exact record of how often the device is switched on (i.e. wear time); the frequency of use of extra features or accessories (such as telecoils and FM systems); as well as the type of listening environment experienced (and the length of time spent within these environments).

The intention of data logging is to provide objective data to help inform patient counselling, as well as to guide any adjustments made to the maps to optimise patient satisfaction and even accelerate improvements in hearing performance (Cochlear Americas, 2013; Cochlear, 2015; Flynn, 2005). CI recipients are encouraged to maximise wear time, to the point where the CI device remains switched on for every waking hour. The rationale is that this exposure and practice is the key to CI success, in terms of speech understanding and language development (NDCS, 2018; AVUK, 2016). Interestingly, however, data logging studies have revealed that the average CI device usage tends to be lower than expected (e.g. Lurie & Lurie, 2014). This “suboptimal” CI device use (in the case of children) has been attributed to factors such as infection or injury; equipment problems; insufficient support (i.e. infrequent clinic visits); and socioeconomic or family stressors (Steacie et al., 2016).

It is noticeable that there has been relatively little investigation into the relationship between wear time and the experience of LE. This potentially reflects a lack of awareness, and even acknowledgment, that using the CI device may continue to be effortful (irrespective of length of exposure). Indeed, reduced wear time may even be a coping strategy for the CI users, in an attempt to manage their LE levels.

It may be the case that speech understanding and language development becomes more proficient (over time) because the CI user may progressively become more skilled in the operation of compensatory mechanisms to fill in the gaps in the impoverished auditory information (provided by the CI device). However, it can be argued that the very fact that these mechanisms are needed in the first place means that there will always be a persistent additional cognitive load for the CI user, even in optimal listening conditions.

Furthermore, it is possible that this potential for extra cognitive load (in quiet) is not necessarily being considered as an ongoing problem. Indeed, it can be argued that the current focus in the support of both children and adults following cochlear implantation primarily targets the challenge of speech perception in noise. For example, in mainstream schools (as well as in further education and higher education facilities), there has been a concerted effort to improve classroom/lecture hall acoustics through acoustic treatment of floors, walls and ceilings, as well as the installation of sound field technology and provision of radio aids (NDCS, 2016; NDCS, 2017; Phonak, 2010).

All of these interventions aim to improve the quality of sound received, irrespective of where the CI user is seated in the classroom or lecture hall (Phonak, 2010). There has also been ongoing research and development of noise reduction algorithms for the CI neuroprosthesis (Loizou et al., 2005; Mauger et al., 2012; Ye et al., 2013; Hersbach et al., 2013). These algorithms are specifically intended to improve the efficiency and efficacy of the CI in speech extraction during challenging listening conditions.

However, despite all these innovations, there has been growing concern about listening fatigue, particularly within the paediatric CI population. For example, surveys of hearing impaired children in mainstream schools have identified parental concerns about cochlear implanted children returning home exhausted at the end of the school day and exhibiting behavioural difficulties during the school day itself, despite these children being apparently well motivated to learn (NDCS, 2016; Damen et al., 2006; Geers & Brenner, 2003).

Furthermore, this fatigue was also still evident in bilaterally implanted children (NDCS, 2016; Steel et al., 2015; Faulkner & Pisoni, 2013). This is in spite of the evidence that bilateral implantation confers considerable advantages (compared to unilateral implantation) in the discrimination of speech in noise (Litovsky et al., 2009, 2012; Aronoff et al., 2010; Chan et al., 2008).

These circumstances suggest that strategies to reduce the impact of challenging listening conditions (i.e. on speech perception) do not reduce the incidence of listening fatigue in CI users.

This fatigue could be argued to arise from excessive LE. However, whilst the link between these two concepts seems plausible, it does need to be borne in mind that this causal relationship has yet to be empirically proven in literature.

The existence of excessive LE is particularly worrying because of the burgeoning research reporting its relationship with increased risk of allostatic injury (caused by excessive LE becoming a stressor: Pichora-Fuller et al., 2016; Pichora-Fuller, 2016 Schneiderman et al., 2005). This, in turn, implicates both mental and physical health problems (particularly via the dysregulation of the HPA axis: Stephens & Wand, 2012).

If listening effort is significantly increased for CI users even in optimal listening conditions, this needs to be addressed in terms of a revision in educational support for deaf children and young adults, as well as a change of focus in CI technological development and rehabilitation. Monitoring for the physical and mental health consequences of LE should also take a much higher priority for all CI recipients. However, none of these changes can be realistically implemented until there is sufficient evidence that electric hearing can cause cognitive duress, even in perfect listening conditions.

This requires the development of a reliable and sensitive clinical test of LE that can accurately measure LE in a wide range of listening conditions. In order for this to be achieved, the methodology of any attempt in this field needs to be carefully scrutinised to determine best practice.

7.7. Limitations in current dual-task design

Throughout the course of the experiments described in this thesis, a series of methodological issues were identified, which potentially limit the effective use of the current dual-task paradigm design. Attempts were systematically made to address these limitations within this current research (which have already been discussed within this chapter). However, additional factors need to be taken into consideration. These are identified and discussed within this section.

7.7.1. CI simulation training

It is difficult to determine the optimum training duration in order to complete perceptual learning and/or adaptation to vocoding, especially when using noise in the background in addition to the spectrally degraded stimuli (Stone & Moore, 2003).

This means that any experiments involving vocoded stimuli need to include an assessment of whether adaptation has truly been achieved before progressing to formal testing.

7.7.2. Adaptive tracking

The presence of individual variation in auditory processing gives rise to the extra consideration that the “starting point” for each participant may differ, in terms of their inherent ability to manage both single and dual-tasking.

However, the use of adaptive tracking to accommodate this (and achieve comparability) does seem to be inconsistent in its efficacy, especially when two types of tracking are implemented together (e.g. the combination of tracking of both channel number and noise level for vocoded stimuli). It may be the case that executing subsequent tracking (i.e. on top of performance levels that had already been tracked) is untenable.

It is also important to recognise that over-performance (relative to the tracked level of performance) is just as problematic as under-performance. This is because over-performance introduces the concern that the participant is not being challenged sufficiently and, therefore, will not be exhibiting representative behaviour.

Under-performance, on the other hand, could indicate that the imposed challenge may have been too great, with the potential consequence of compromised cognitive processing.

Therefore, whenever tracking procedures are implemented, care needs to be taken in ensuring that the intended performance level is actually yielded in all participants tested.

7.7.3. Task difficulty

In order for the dual-task paradigm to execute correctly, there has to be appropriate prioritisation of cognitive resources between the primary and secondary tasks (with the primary task always taking precedence). In addition, the combination of demand from both the primary and secondary tasks must not completely exhaust cognitive resources.

The experience from this current research suggests that that memory tasks are particularly taxing in terms of cognitive resources. This is undesirable because there is an increased risk of excessively depleting cognitive capacity and invalidating the measurement. In addition, stimulus presentation order within the dual-task paradigm could inadvertently lead to inappropriate diversion of cognitive resources away from the primary task. This could potentially be due to the stimulus salience of the secondary task increasing, thus attracting more attention (and cognitive resources). Therefore, ideally, stimulus presentation for both the primary and secondary tasks should be simultaneous.

Also, even when a task appears to be simple, it can unintentionally incur other cognitive processes which impinge on the residual cognitive capacity. For example, the counting of visual targets should (in theory) only tax the visual sensory modality. However, it is tenable that the process of counting could additionally be recruiting verbal rehearsal (and other forms of language processing). This could then encroach into available cognitive capacity needed, and deplete it inappropriately.

Therefore, there needs to be caution when selecting the primary and secondary tasks for the dual-task paradigm, to ensure appropriate resource allocation so that the resulting measurement is meaningful.

7.7.4. Primary task stability

In order for the dual-task paradigm to function as designed, another aspect to monitor is whether there is primary task stability. It is essential that primary task performance remains constant, irrespective of whether the secondary task performance is concurrently happening or not. This stability can only be maintained if the primary task is being prioritised in terms of cognitive resources (which is fundamental to the dual-task paradigm). Therefore, if any deterioration in primary task performance occurs, when the participant is dual-tasking, there should immediately be concern that the secondary task has diverted cognitive resources away from the primary task.

However, this monitoring of primary task stability has the complication that it is often difficult to determine exactly what extent of deterioration in primary task performance actually constitutes instability in primary task performance.

In the case of the present research, statistical analyses were executed to detect significant decrements in primary task performance and, therefore, primary task instability. However, this approach has two limitations. Whenever an absence of effect is detected, it cannot be automatically assumed that primary task stability has been achieved. Furthermore, if a statistically significant effect is yielded (i.e. there appears to be instability), it does not offer a scalar quantity related to the level of instability being detected.

Therefore, establishing criteria that enable the definition of the exact point at which primary task stability becomes primary task instability is critical for confident implementation of the dual-task paradigm.

7.7.5. Perceptual response

To maximise sensitivity of the measurement (within the dual-task paradigm), the true nature of the participant's response needs to be understood at the perceptual level. The brain can be conceptualised to be not just a passive receiver of the information, but instead an active decision maker (Green & Swets, 1966). Accordingly, a perceptual response falls into four categories: a "hit", a "miss", a "correct rejection" and a "false alarm". In the case of human perception, the central nervous system is "noisy". This is attributed to neurons being constantly active and sending impulses to the brain, even in the absence of stimuli (Green & Swets, 1966; Heeger, 1997). This has the implication that human perception can make errors regarding the category of perceptual response.

By differentiating between these different perceptual responses, it becomes possible to understand a participant's accuracy in terms of both sensitivity and response bias. This is because there is an ongoing "trade-off" between hit and false alarm rates, whenever a perceptual decision is made (Green & Swets, 1966; Heeger, 1997). By being more liberal in one's response criterion, hit rates can be increased (but at the cost of increased likelihood of false alarms). On the other hand, if a response criterion is more conservative, false alarm incidence is reduced, (but at the cost of missing the target more frequently).

A consistent phenomenon in this present research was over-reporting by the participants. This was unexpected because the original hypothesis postulated that, as task difficulty increased (when dual-tasking), the participants' secondary task performance would worsen, i.e. more targets will be missed (leading to under-reporting).

It is possible that over-reporting was symptomatic of a relatively liberal response criterion (i.e. hits were reported more often than not, increasing the incidence of false alarms). However, this cannot be concluded definitively, because the nature of the current digit stream measure cannot discriminate between the different types of perceptual responses. For example, over-reporting could reflect a misidentified digit, or alternatively could constitute a false alarm.

This uncertainty (regarding nature of perceptual response) also extends to the participant's other answers within the secondary task. For instance, it is not known whether a successfully detected target is indeed a genuine detection (versus a false alarm, or a misidentified digit). Likewise, when the target is missed, it is not known whether it is a genuine "miss".

To further complicate matters, over-reporting may arise from the participant simply guessing what the answer is. Also, the participant may even be self-monitoring how difficult they are finding the secondary task (and the dual-task in general), and may be adjusting their answers as part of their attempt to improve accuracy (i.e. the harder they find the task, the more they may presume they have missed targets and, therefore, compensate in their answer). Inevitably, using these strategies, the participants' answers may not reflect their actual perceptual experience (and the data obtained might consequently be compromised).

7.7. Limitations in design

One approach to enable this much-needed differentiation in perceptual response is to apply signal detection theory (SDT). SDT is a statistical technique designed to quantify the ability of the system to discriminate a signal against a background of noise (Green & Swets, 1966). A behavioural task enabling this application is the two-alternative forced choice (2afc) detection task.

The 2afc involves presenting the participant with two options (Blackwell, 1953; Luce, 1963, 1993; Fechner, 1860; Ulrich & Miller, 2004). Only one contains the target stimulus and sometimes the stimulus is completely absent. Both the options can be presented concurrently (within two spatial windows), or sequentially in two intervals (this temporal derivative is sometimes called the two-interval forced choice: Fechner, 1860).

When applying this 2afc technique within the dual-task paradigm, the target stimuli could be a visual grating. This is a type of stimulus extensively used in visual perception research. Gratings are fundamentally created from oscillations in luminance over space, where spatial frequency, contrast, phase and orientation can be tailored to suit the experimental design (Campbell & Robson, 1968). Accordingly, in a given trial within the dual-task, the participant would not only need to recall the sentence they heard, they would also be required to state which of the two windows contained the grating. To adjust the difficulty level of this new secondary visual task, the visual grating could vary in either intensity, and/or duration.

Ultimately, any uncertainty as to the nature of the perceptual responses undermines the evaluation of whether resource allocation within the dual-task paradigm is appropriate. Accordingly, there is a need to clearly identify each category of perceptual response. Experimental design incorporating signal detection theory appears to be an appropriate way forward in this regard.

7.7.6. Scaling LE

Whilst statistically significant changes in secondary task performance are being detected when dual-tasking in different listening conditions (thus suggesting increases in LE), these different percentages unfortunately do not immediately lend a scalar quantity, or unit, for LE measurement.

Thus, it is currently not known what level of LE has actually been incurred if secondary task accuracy deteriorates by a difference in 1%, let alone some other percentage. For instance, it may be the case that substantial LE has to be experienced before a particularly large decrement in secondary task accuracy is detected.

This creates considerable difficulties regarding the elucidation of the impact that LE has on the individual. In particular, it has yet to be understood what level of LE begins to trigger health problems. Future experiments, therefore, need to be able to generate measurements that can be sufficiently graduated to constitute a scale. Only then will it be possible to begin to assess the cognitive burden a particular position on the LE scale imposes on the individual.

7.7.7. Subjective ratings

Within this current research, subjective ratings did not always corroborate the behavioural results. This may be due to inherent weaknesses and, therefore, insensitivity within the behavioural measurement. However, it cannot be ignored that subjective ratings have their own limitations that need to be factored in.

A particular concern is that subjective ratings can show variability, not only between individuals (i.e. inter-individual) but also within individuals (i.e. intra-individual). This gives rise to the concern that subjective ratings may not be consistently reliable, especially if used on a repeated basis. There is also the question of whether, or not, the participant is actually defining perceived effort appropriately (i.e. as the researcher intended). For example, the participant may be answering in accordance to how much effort they believe they needed to apply in order to succeed, rather than their actual subjective experience of LE. Another potential source of confusion is the participant basing their subjective ratings on their fatigue level, which is not necessarily synonymous with LE.

Nevertheless, subjective ratings are a relatively intuitive and easy tool to gather additional data about LE, which can then provide some limited across-methods triangulation. This triangulation can provide a valuable perspective as to the likelihood that resource allocation within the dual-task paradigm is appropriate.

7.7.8. Power

The difficulty with trying to develop a test, that is intended to become a “gold standard” clinical procedure, is that an appropriately high level of power must be achieved. This becomes especially difficult when this test has to be applied to CI users, who are recognised to be a highly heterogeneous population (Blamey et al., 2013; Green et al., 2007; Holden et al., 2013).

In theory, experimental design can be adjusted such that there are fewer variables to control, which may potentially necessitate smaller, and therefore more feasible, sample sizes for recruitment and testing.

However, such fine-tuning of experimental design is difficult when there is little empirical data already available to guide the decision as to which variables to disregard and which to control. Indeed, within this current research, it was this lack of prior data that prevented achievement of power. This issue applied even in experiments involving normal hearing participants, where the sample size (necessitated by experiment design) exceeded 60.

Ultimately, a clinical test needs to be inherently robust enough such that it can be applied to the individual patient without the use of group norms. This is a challenging criterion to fulfil, and may require considerably more collection of baseline data, as well as refinement of both experiment and test design, before clinical assessment of LE can become a reality.

7.8. Clinical feasibility of LE testing

The original question, which this research attempts to address, is whether the dual-task paradigm itself is a feasible framework for the development of a clinical test of LE. The evidence from the present research appears to suggest that this behavioural approach does seem capable of detecting changes in LE and, therefore, there does appear to be clinical potential for this dual-task design.

However, the digit stream version of this dual-task design is currently achieving very low power. In considering the implications of this low power, it is necessary to recognise that an integral part of power analysis is effect size. Effect size is concerned with the strength of the relationships among research variables. For effect size to be large, it is necessary for the researcher to collect evidence to demonstrate that the dependent and independent variables are strongly inter-related. This will then contribute to increasing power (Cohen et al., 2003).

Therefore, a potential avenue to improving power in the current dual-task design is to enhance the sensitivity of the measurement of the dependent variable, so that its relationship to the independent variables can be more accurately assessed. In this case, the dependent variable is visual accuracy (currently designed to index LE).

The independent variables include factors such as the presence of noise (and its level) and the nature of the speech, whether normal or spectrally degraded (i.e. by the noise-excited vocoding, or the inherent limitations of the CI device).

Bearing in mind the potentially constrained cognitive capacity of the CI user in quiet (as discussed earlier in this chapter), there may only be small gradations in the re-allocation of the remaining cognitive resources as listening conditions become more challenging (e.g. decreasing SNR). Thus, the granularity of the measurement (of the dependent variable) needs to be high to enable sufficient sensitivity. The 2afc methodology potentially offers this granularity and, in turn, sensitivity. This would allow the necessary exploration and assessment of slight changes in dual-task performance as the CI user approaches the brink of cognitive overload.

This improved sensitivity could also contribute to the evidence regarding effect size, potentially enabling the higher level of power necessary for a feasible clinical test of LE that can be applied to individual listeners.

However, within the clinical setting, there is still the issue of restricted time available for an audiological appointment. Therefore, the incorporation of any additional testing in the form of LE assessment becomes a challenge. It has already been argued in this thesis that the dual-task approach could enable simultaneous testing of both speech intelligibility (i.e. primary task performance) and LE (i.e. inferred from secondary task performance).

The time saving from this “two-for-one” approach could be further improved by reducing the number of test conditions required by the LE assessment. Indeed, adaptive tracking of SNR level (currently used to define level of acoustic challenge) is potentially only needed in the developmental stages of the LE test. It may transpire that further methodological refinement could lead to its removal altogether. At this point, fixed noise levels could be applied instead. This is particularly appropriate as these are already implemented in commonly used speech intelligibility testing (which would again reduce testing time).

However, it does need to be recognised that any dual-task paradigm does require baseline data related to the performance of the secondary task alone, before its combination with the speech intelligibility testing. This is because this baseline is critical in the calculation of LE in quiet listening conditions.

This gathering of baseline data would inevitably take valuable time (within the audiological appointment). A pragmatic solution is therefore needed, and this could be the exclusive testing of LE in quiet. In light of the findings in this present research, it can be argued that the level of LE in quiet could be extremely informative by itself. In particular, clinically there is a need to understand the “starting point” at which the CI user is performing (i.e. in the best possible listening condition). Indeed, if LE is already high in this ideal state, auditory processing can only be further compromised by additional demands present in the typically noisy world of sound. If this “starting point” could be optimised (e.g. by tuning the map), then this has the potential to aid efficiency of speech understanding in more adverse listening conditions.

Measurement of LE, therefore, could potentially be implemented with relatively little additional demand on clinical time, while providing invaluable information enabling the audiologist to implement the most efficient and efficacious programming strategy to optimise cognitive capacity for auditory processing. In addition, objective LE assessment could enable evidence-based patient counselling related to healthy management of their LE levels. This is especially important in light of the risk of allostatic injury (i.e. from the chronic stress response) which may be incurred from persistent high levels of LE.

7.9. Conclusion on feasibility of dual-task paradigm

In conclusion, the results of this thesis suggest that the dual-task paradigm could feasibly form the framework for a clinical test of LE in the CI user population. However, there is a crucial proviso: any researchers using the dual-task paradigm should consider the tangible possibility that, for CI users, the cognitive toll of listening in quiet is already profound. Indeed, this cognitive burden could be so considerable that there is extremely limited capacity left to process any further task-related demand. Therefore, any secondary task will need to enable sensitive and accurate detection of very small changes in resource allocation in order to reliably measure LE in the CI user. The ongoing search for such a refined secondary task is particularly worthwhile because, ultimately, understanding the impact of LE could be the key to improving the quality of life for hundreds of thousands of CI users worldwide, today and in the future.

APPENDIX

Appendix 1.1: Protocol guidelines for a CI audiological appointment (T. Twomey, Consultant Clinical Scientist of Auditory Implant Programme, personal communication, November 28, 2014)

CI audiological (sometimes called “programming” or “tuning”) appointments include the following:

- Discussion regarding performance, progress, concerns and technical issues
- Otoscopy and examination of implant site(s)
- Counselling regarding appointment procedure and expectations
- Implant assessment and programming of map (i.e. allocation of electrodes to sound frequency), including:
 - Electrical threshold measurement
 - Evaluation of electrode performance
 - Measurement of loudness growth by electrically evoked stapelial reflexes, and/or behavioural techniques (such as loudness scaling and observation).
- Other electrophysiological measurements as required, such as:
 - Neural Response Telemetry (NRT)
 - Neural Response Imaging (NRI)
 - Auditory Nerve Response Telemetry (ART)
- Evaluation of mapping in “live” mode for audibility and loudness discomfort at each increase in mapping level for each implant (unilaterally and binaurally if applicable).
- Outcome measures as appropriate, such as:
 - Aided thresholds (in accordance to the British Society of Audiology’s recommended procedures)
 - Auditory Speech Sound Evaluation (ASSE) Test
 - Arthur Boothroyd Word Test
 - Bench, Kowel, Bamford (BKB) Sentence Test (Standard and/or Adaptive versions)
 - City University of New York (CUNY) Sentence Test
 - McCormick Toy Test (Live Voice and/or Automated versions)
- Appropriate creation and saving of maps (including progressive maps if required)
- Discussion with patients and/or parents regarding the most suitable settings available on CI device, such as telecoil (taking in account of individual needs and/or data logging)
- Patient and/or family debrief
- Appropriate recording and reporting of appointment proceedings.

Appendix 2.1: Table of the protocol used for the additional testing session required for the Pilot Study to enable neurocognitive and speech-in-noise assessment, with detail regarding whether training was involved, as well as speech type (if applicable), noise type (if applicable) and total number of trial (or total duration). Note that the same protocol is used for CI users with only the exclusion of performing the CCRM in the vocoded condition. Red numbers with asterisks indicate the steps in the protocol where counter-balancing and randomisation occurred.

Step	Training, Practice or Test?	Task Type	Speech Type	Noise Type	Total number of trials/ Total duration
1 *	Test	CCRM	Normal	Speech-shaped noise at variable SNR	25 trials maximum (repeated three times in a row)
		CCRM	Vocoded (6 channels)	Speech-shaped noise at variable SNR	25 trials maximum (repeated three times in a row)
2 *	Training	CANTAB	N/A	N/A	30 minutes
	Practice & Test	CANTAB sub-test: Motor Screening	N/A	N/A	2 minutes
	Practice & Test	CANTAB sub-test: Reaction Time Index	N/A	N/A	5 minutes
	Practice & Test	CANTAB sub-test: Paired Associates Learning	N/A	N/A	10 minutes
	Practice & Test	CANTAB sub-test: Spatial Working Memory	N/A	N/A	8 minutes
	Practice & Test	CANTAB sub-test: Rapid Visual Processing	N/A	N/A	10 minutes
	Practice & Test	CANTAB sub-test: Delayed Matching to Sample	N/A	N/A	7 minutes
3 *	Practice & Test	Reading Span	N/A	N/A	10 minutes
4 *	Practice & Test	WASI sub-test: Block Design	N/A	N/A	5 minutes
	Practice & Test	WASI sub-test: Matrix Reasoning	N/A	N/A	10 minutes

Appendix 2.2: Detail about the CANTAB (Cambridge Neuropsychological Automated Test Battery) implemented in Pilot Study.

The Motor Screening Task served as a warm-up induction task (as well as a check of sensorimotor function) and entails a series of individual pink crosses appearing at random locations on a black background. Participants needed to press on this cross (with the index finger of their dominant hand) when the cross turned green. Data was not formally collected for this test.

The Reaction Time Index (RTI) test enabled testing of motor and mental response speed, as well as impulsivity. Participants were presented with a semi-circular array of five rings and a “press-pad”. During the course of this test, participants were required to keep pressing the press-pad with their index finger (of their dominant hand) until a yellow dot flashes at random in one of the five rings. At this point, participants were required to move their finger off the press-pad and press (with the same index finger) inside the ring they saw the yellow dot appear in. The outcome measure of movement time (i.e. the time taken for the participant to correctly press the ring where they had just seen the yellow dot appear in) was focused on for further analysis.

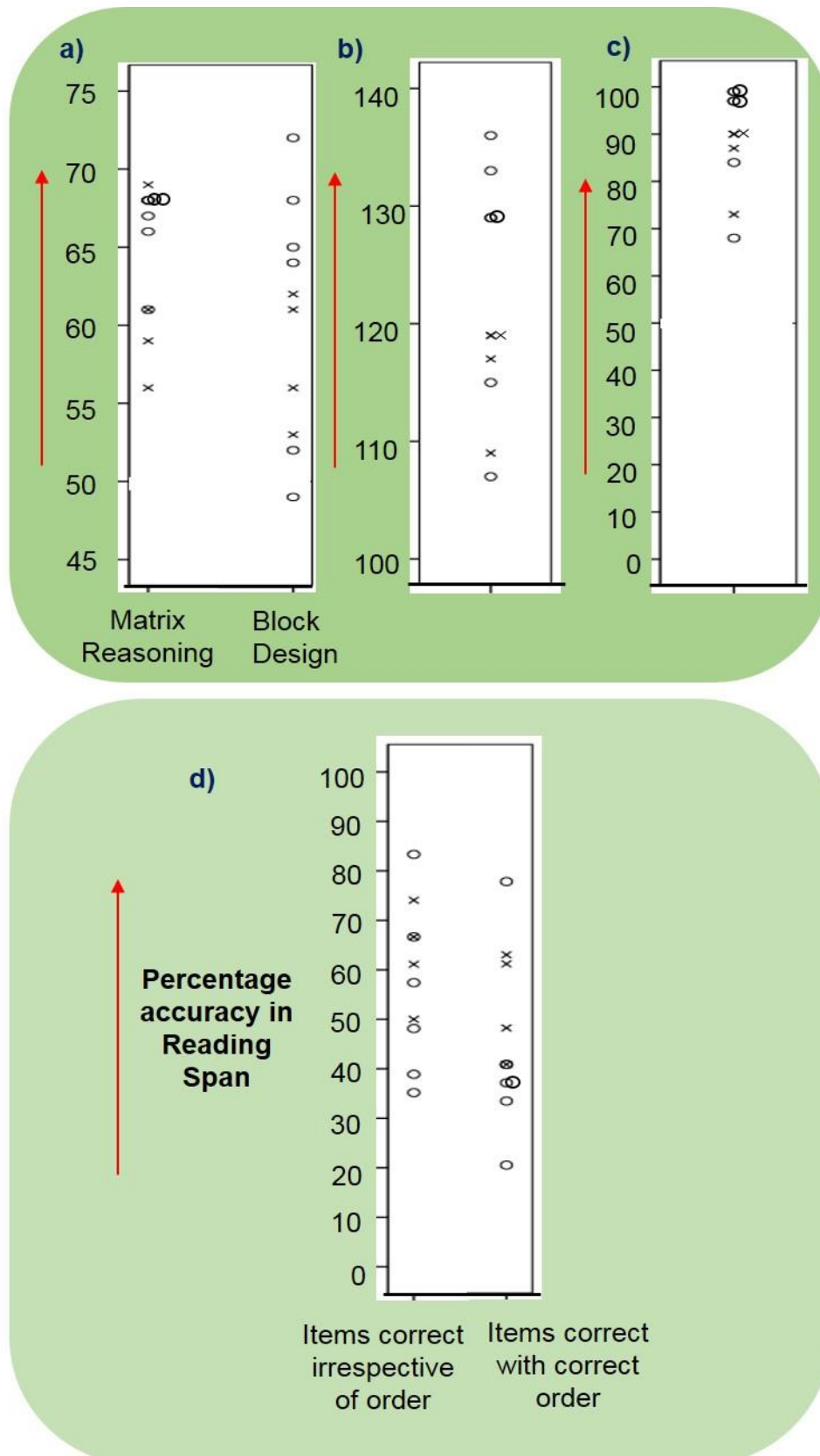
The Paired Associates Learning (PAL) test assessed episodic visual memory and new learning. Boxes were displayed on the screen and were opened in randomised order. One or more of these boxes contained a pattern (which consisted of one or more colours). Participants were required to remember the location of these patterns and upon probing (with the presentation of the target pattern at the centre of the screen), participants touched the box where they believed that pattern was originally located. The difficulty level of this memory task increased throughout the test, going from 2 patterns, to 4 patterns, 6 patterns and finally 8 patterns. The following outcome measures were chosen by this study for further analysis: total number of errors made throughout entire task; and the number of correct choices that were made within the first attempt of a pattern array.

The Spatial Working Memory (SWM) test required the retention and manipulation of visuospatial information. The test began with a number of coloured squares presented on the screen. The aim of the test was that, by touching these boxes and using a process of elimination, the participant should find a token in one of these boxes, and use these tokens to fill up an empty column on the right hand side of the screen. The number of boxes was gradually increased (from 4 and 6 boxes) until it became necessary to search a total of 8 boxes. The colour and position of these boxes were changed from trial to trial in an attempt to discourage the use of stereotyped search strategies. The following outcome measure was utilised for this study: total number of errors; and strategy (i.e. the number of distinct boxes searched, with no revisiting or error).

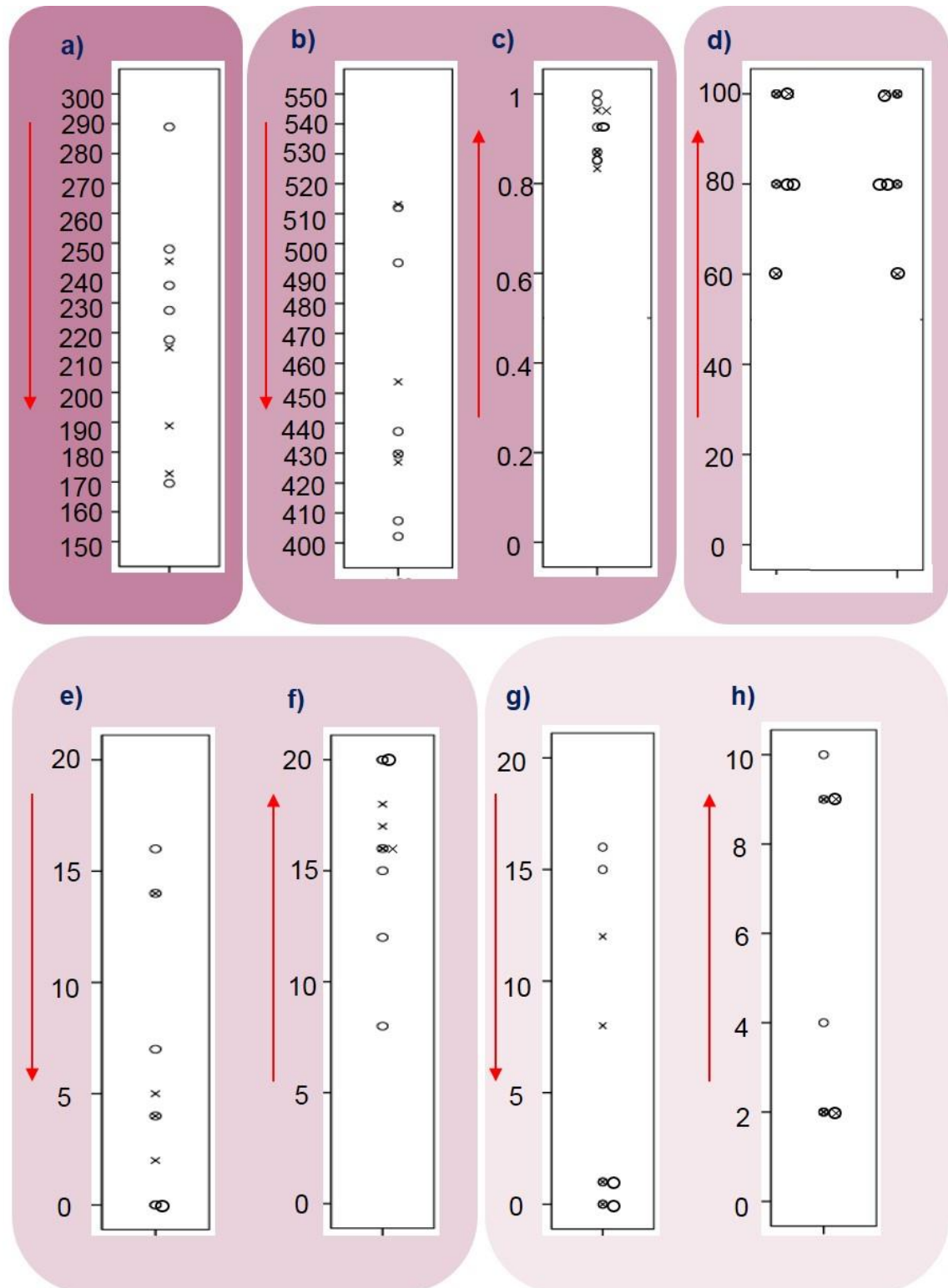
The Rapid Visual Information Processing (RVP) test was a measure of sustained attention. During the test, a white box appeared in the centre of the computer screen, inside which digits (from 2-9) appeared in a pseudorandom order at the rate of 100 digits per minute. Participants were required to detect target sequences of digits (i.e. 2-4-6, 3-5-7 and 4-6-8) and to register their response by pressing a press-pad upon completion of the presentation of the final number of the detected sequence. At first, participants only had to detect one sequence (i.e. 2-4-6) but then quickly graduated to needing to detect all three sequences. The chosen outcome measures chosen for the RVP were mean latency of correct responses and the probability of hits.

The final test, the Delayed Matching to Sample (DMS) test, assessed forced-choice recognition memory for non-verbal patterns, testing both simultaneous matching (a check of perceptual ability) and short-term visual memory. Participants were shown a complex visual pattern that was multi-coloured and multi-textured. Then, after either a brief delay of 4 or 12 seconds, or no delay (with the target pattern remaining visible for the duration of the response period), the participant must touch the pattern (from a choice of four answers) that exactly matched what they had just been presented. Overall percentage correct in recall after a 4-second delay and a 10-second delay were the chosen outcome measures.

Appendix 2.3: Dot plots of performances obtained the WASI and Reading Span test for all participants (Pilot Study). **(a)** t-scores of all participants in the performance sub-tests of the WASI, i.e. the Matrix Reasoning and Block Design. **(b)** Non-verbal IQ as calculated from performance sub-tests. **(c)** Percentile of non-verbal IQ. **(d)** Percentage accuracy achieved in Reading Span irrespective of order and with items recalled in order. O = NH control (in normal listening condition), X = CI user. Red arrow indicates superiority of score.

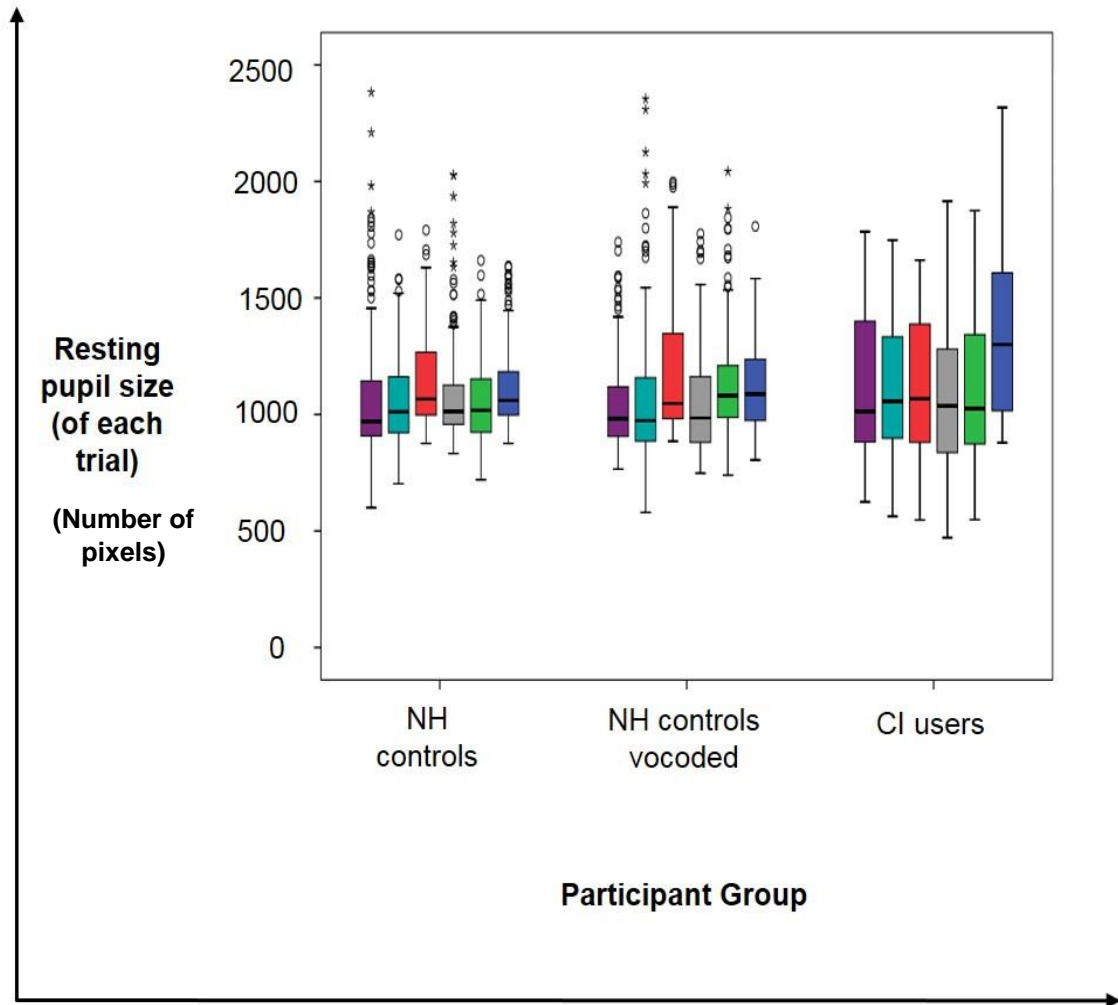


Appendix 2.4: Dot plots of performances obtained in the five sub-tests of the CANTAB for all participants (Pilot Study). **(a)** Reaction Time Index: mean movement time (in milliseconds). **(b)** Rapid Visual Processing: mean latency of response (in milliseconds). **(c)** Rapid Visual Processing: probability of hits. **(d)** Delayed Matching to Sample: percentage correct in recall after a 4-second or 10-second delay. **(e)** Paired Associates Learning: total number of errors. **(f)** Paired Associates Learning: Total number of correct responses given on first attempt. **(g)** Spatial Working Memory: total number of errors. **(h)** Spatial Working Memory: Number of correct boxes searched without revisiting. O = NH control (in normal listening condition), X = CI user. Red arrow indicates superiority of score.



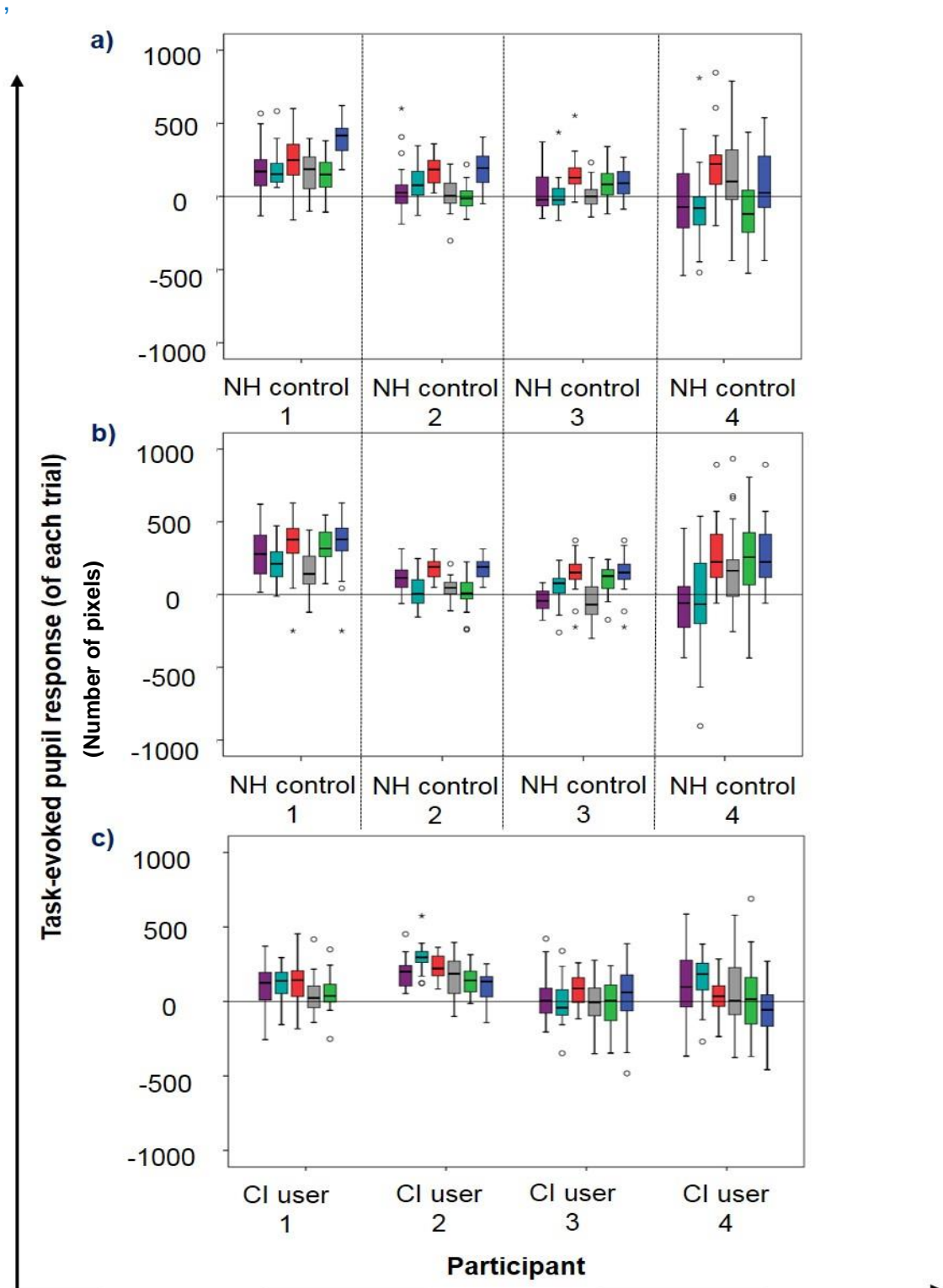
Appendix 2.5: Box plots showing variability of median pupil size obtained at rest (during measurement at baseline) of all trials within each test condition for all NH controls (in normal and vocoded listening conditions) and all CI users (Pilot Study).

Purple = Visual task only in quiet, Turquoise = Auditory task only in quiet, Red = Dual-tasking in quiet, Grey = Visual task only in noise, Green = Auditory task only in noise, Blue = Dual-tasking in noise. Speech-shaped noise (when applicable) was at -5dB SNR for CI users and NH controls in normal listening conditions, and 0dB for NH controls in vocoded listening conditions.



Appendix 2.6: Box plots showing variability of task-evoked pupil response obtained in all trials within each test condition for participant (Pilot Study). Baseline correction was applied in all trials (i.e. the subtraction of median pupil size in baseline epoch of corresponding trial). **(a)** NH controls being tested in normal listening conditions. **(b)** NH controls being tested in vocoded listening conditions. **(c)** CI users.

Purple = Visual task only in quiet, Turquoise = Auditory task only in quiet, Red = Dual-tasking in quiet, Grey = Visual task only in noise, Green = Auditory task only in noise, Blue = Dual-tasking in noise. Speech-shaped noise (when applicable) was at -5dB SNR for CI users and NH controls in normal listening conditions, and 0dB for NH controls in vocoded listening conditions.



Appendix 3.1: Explanation regarding choice of visual accuracy scoring method for Experiments 1-8.

The results from Experiments 1-8 revealed three different types of participant response related to the secondary visual task:

1. Correct response (i.e. the participant reported the exact number of targets presented)
2. Under-reporting (i.e. the participant reported seeing fewer targets than actually presented)
3. Over-reporting (i.e. the participant reported seeing more targets than actually presented)

As part of the development of the visual task, the digit stream speed had been intentionally set at sufficient difficulty to reduce the incidence of correct responses. The original hypotheses were that under-reporting of visual targets would occur in dual-tasking and that this under-reporting would also increase as the difficulty of the dual-task increased.

Over-reporting had not been predicted within the original hypotheses, and its actual incidence proved to be low. For example, in the key experiments (7 and 8) involving CI simulation and CI users respectively, the percentage of over-reporting for all trials, across all 55 participants (30 NH controls and 25 CI users), for all test conditions ranged from only 4.1% to 7.8%.

Nevertheless, in the preliminary analysis of the data, it was necessary to clarify whether under-reporting had the same status as over-reporting, in terms of constituting error. This meant evaluating whether one reporting behaviour was more important than the other, as a determinant of the difficulty of the dual-task. This evaluation initially involved the application of binary regression models to each of participant response for both key experiments (7 and 8).

To render the data binary, the participant's response in each trial were parsed according to the following criteria:

- Binary regression model for over-reporting:
 - Value = 1 if absolute difference between participant's response and actual number of targets was greater than zero.
 - Value = 0 if otherwise.
- Binary regression model of under-reporting:
 - Value = 1 if absolute difference between participant's response and actual number of targets was less than zero.
 - Value = 0 if otherwise.
- Binary regression model for correct responses:
 - Value = 1 if absolute difference between participant's response and actual number of targets was zero.
 - Value = 0 if otherwise.

In all of these binary regression models, predictors included test condition; target presentation frequency (i.e. the actual number of targets presented in the given trial); and trial position within the total sequence (e.g. the 10th trial in the sequence comprising 18 trials).

Overall, for the binary regression models of the under-reporting for both experiments, there was a significant main effect of test condition ($\chi^2 (3) = 7.82, p=.005$) and also of target presentation frequency ($\chi^2 (5) = 131.3, p<.001$). Post-hoc pairwise comparisons (Bonferroni-corrected) were pursued, revealing that as test condition became more difficult, under-reporting increased ($p=.005$). In addition, it was found that as the number of targets presented (within the given trial) increased, under-reporting also increased ($p<.001$).

Similar main effects were found for the binary regression models of correct responses for both experiments (test condition: $\chi^2 (3) = 2.4, p=.04$; target presentation frequency: $\chi^2 (5) = 116.4, p<.001$). Post-hoc pairwise comparisons (Bonferroni-corrected) revealed that correct responses decreased as test condition increased in difficulty ($p=.04$) and number of targets presented increased ($p<.001$).

Notably, the binary regression models for over-reporting were unable to converge. This lack of convergence was due to the relative sparseness of trials where over-reporting actually occurred.

Under-reporting, therefore, showed the systematic pattern predicted in the original hypotheses. In contrast, over-reporting was not occurring sufficiently for any such relationship, or predictive pattern, to be detected. Thus, under-reporting appears to have a higher status as an indicator of visual error (and, therefore, listening effort) than over-reporting.

In the light of this information, different scoring methods for visual accuracy were considered. It was necessary to find the analysis method that best suited the nature of the data collected, particularly in the key experiments (7 and 8).

In all trials of these experiments, the total number of targets presented was always the same. Across the 18 trials in each test condition, there would be 3 iterations of each possible target presentation frequency (6, 5, 4, 3, 2 or 1 target in a given trial). The presentation of these iterations were randomised in each test condition.

After some deliberation, it was decided that there were 3 potential scoring methods that might be appropriate:

1. Percentage of trials with correct response

The strength of this scoring method is that it is a simple and unbiased technique of assessing visual performance (by extracting just the correct responses). However, the exclusion of the under-reporting data may be to the detriment of the sensitivity of this visual outcome measure. This is because the binary regression models have revealed that the occurrence of under-reporting has a systematic pattern and relationship with the test condition. Therefore, the under-reporting behaviour appears to be of informative value with regards to elucidating LE.

2. Total number of visual errors

This involves the calculation of the absolute difference between the correct answer and the participant's actual answer for each trial, and then summation across all trials in each test condition. This technique is unbiased, ensuring equal penalty for both under and over-reporting (i.e. if the target presentation frequency was 4, a response of "2" or "6" would produce an equal error score of 2 points). Accordingly, a low error count should indicate high accuracy.

However, the application of this scoring method is not necessarily as straightforward as it first appears. This is because it is possible that the same number of errors can be yielded, by different participants, via diverse patterns of responding.

For example, two participants could display the same error score of 3. Yet, in the case of the first participant, this error score could arise from a mistaken response of “1” being given in a trial where there was actually 4 targets presented. While, in the case of the second participant, their error score of 3 could be produced from the mistaken response of “1” being given in three trials, in which (on all occasions) there were 2 targets presented.

Thus, the first participant was mistaken on only one occasion, whilst the second participant was repeatedly mistaken. It could be argued that the second participant is actually less accurate than the first participant (because of this repeated erroneous behaviour) and yet this is not being represented within the total number of errors.

Therefore, it cannot be confidently assumed (using this method) that a low error count always reflects high accuracy. This, in turn, ultimately impacts on the resulting validity and reliability of the LE measurement, if this scoring method was implemented.

3. *Percentage of targets detected*

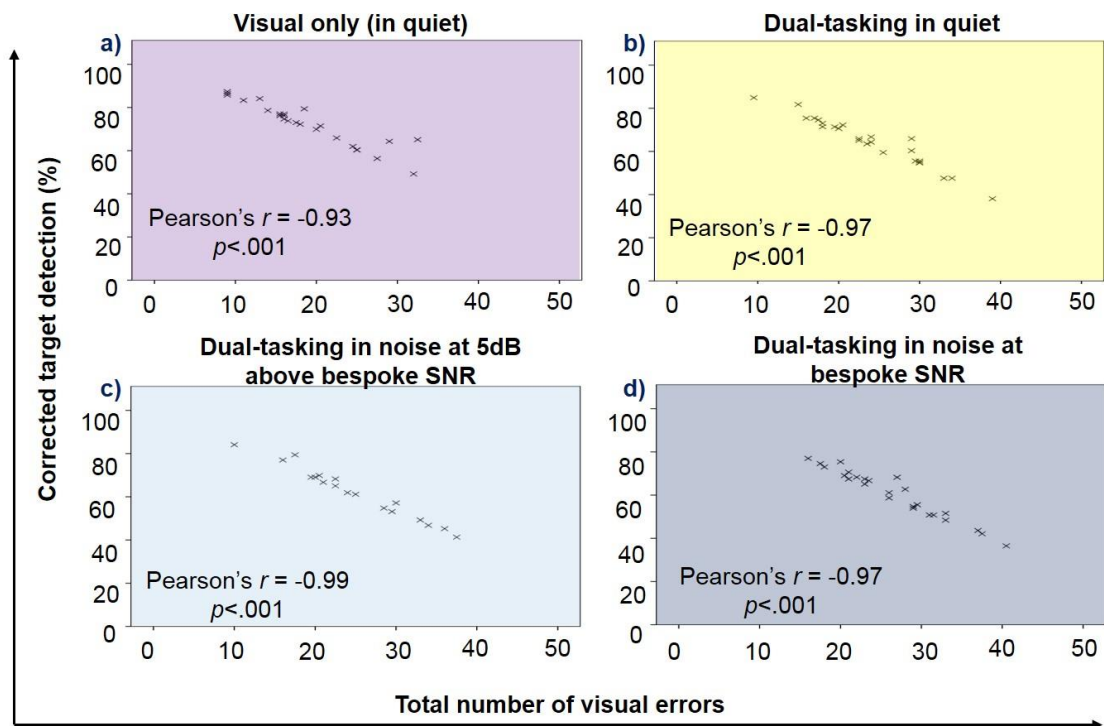
This scoring method summates the targets the participant reports seeing across all trials. In this process, any instances of over-reporting are corrected back to the actual target presentation frequency. The final value is calculated as a percentage of the total number of targets actually presented in the entire test condition. The correction of the over-reporting means that no percentage can exceed 100%, enabling a better sense of proportion and, therefore, relativity, than a score in excess of 100% where the end point of the scale cannot be determined.

This approach to measuring visual accuracy has strength in that it integrates the under-reporting behaviour, as well as the correct responses, into the score. A potential weakness, however, is that the correction (and, therefore, removal) of over-reporting could be a source of bias.

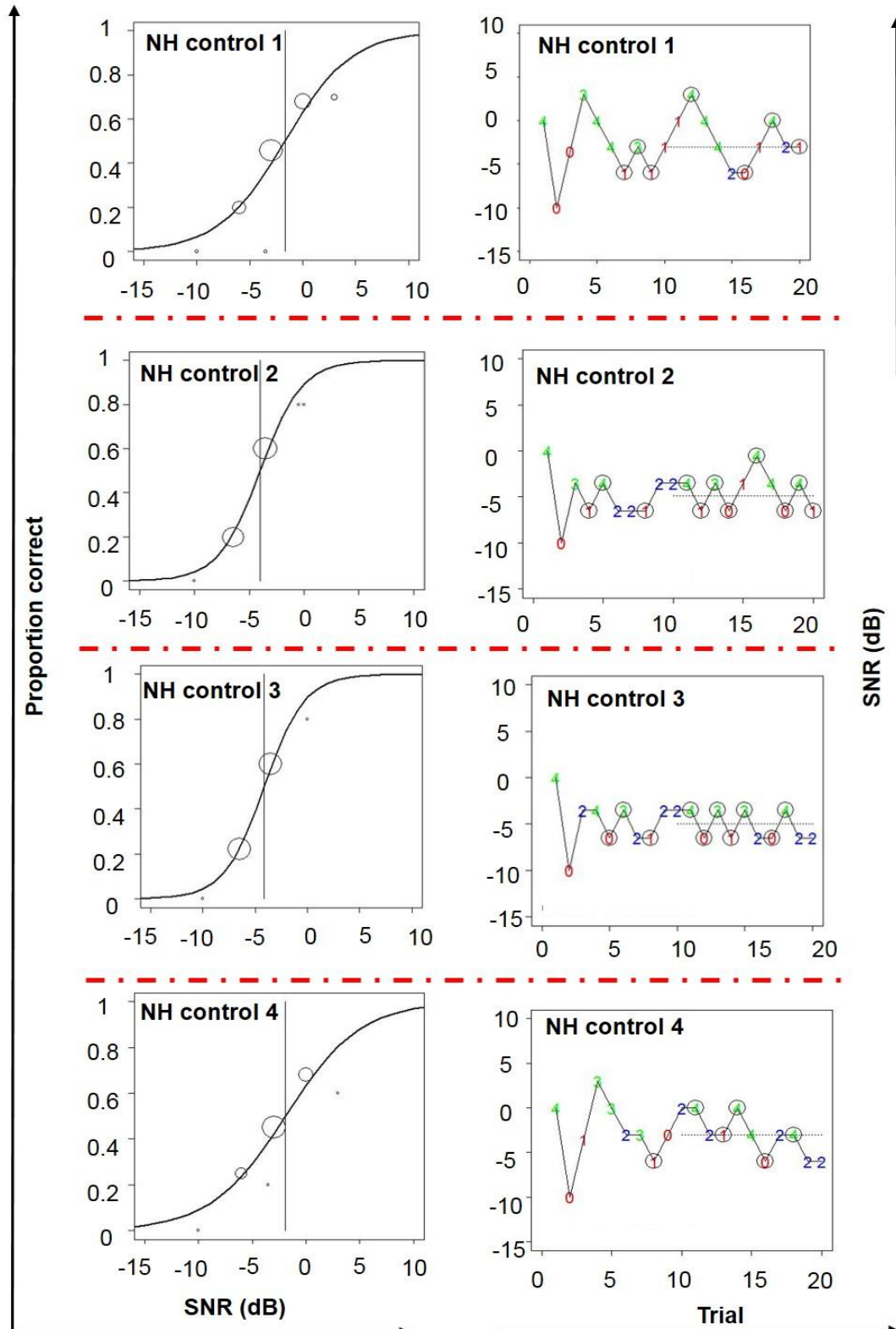
The binary regression models indicated that it was under-reporting that was potentially the most informative measure, because over-reporting did not occur sufficiently frequently to interact in any systematic way. However, although this over-reporting behaviour is sparse, this still does not change the fact that exclusion of over-reporting could potentially change the nature of the data trend. One way to evaluate whether this has occurred is to calculate the correlation between the scoring methods of percentage of targets detected (which removes over-reporting) and total number of visual errors (that encompasses both over and under-reporting).

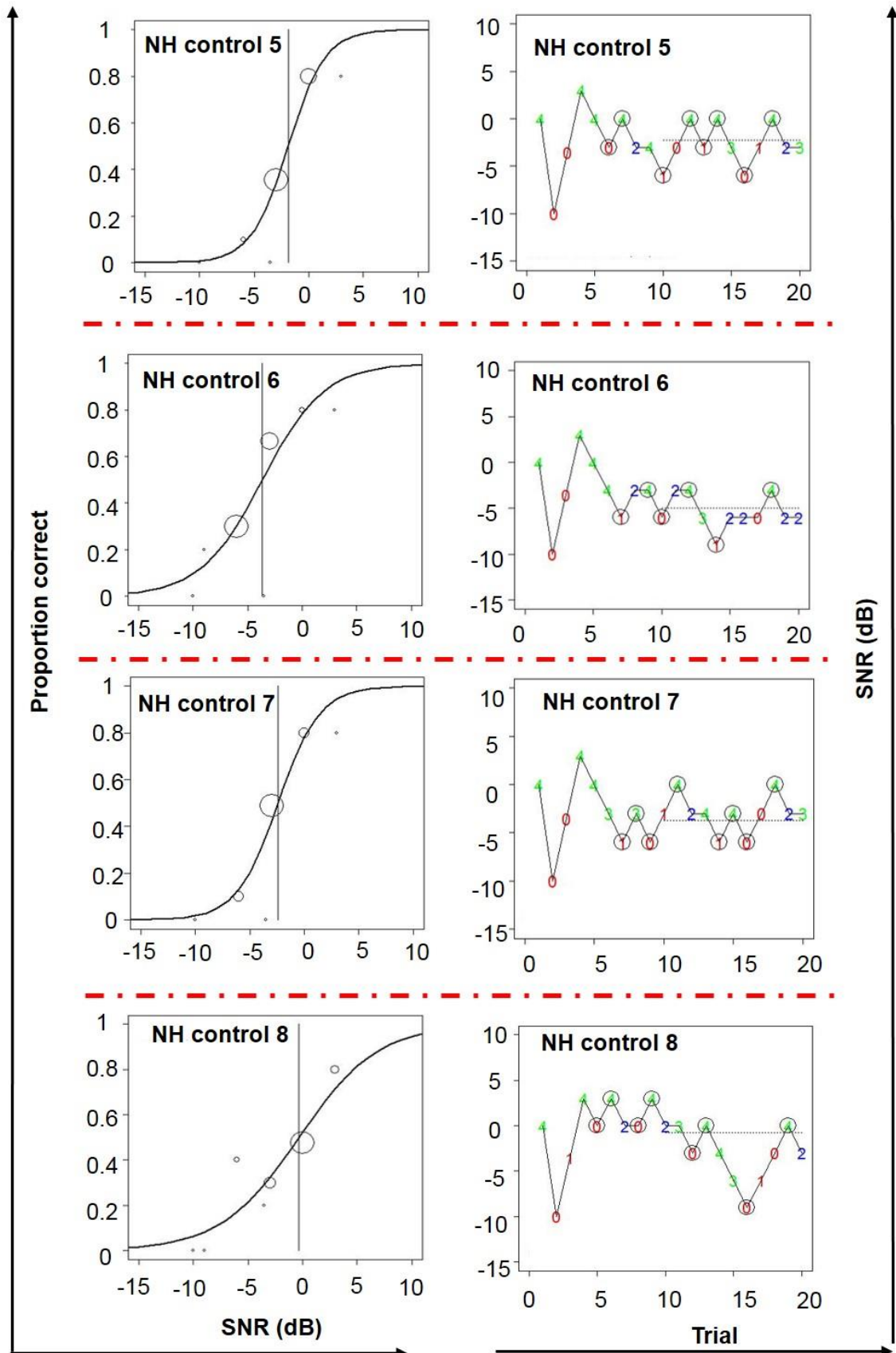
The results from Experiment 8, shown on next page, demonstrated that not only is the relationship between these two visual outcome measures significant, they are also strong. This provides some reassurance that the impact of the over-reporting correction appears to be minimal. This, in turn, means that this correction is potentially of little hindrance in achieving both reliability and validity in the visual accuracy score and, therefore, LE measurement.

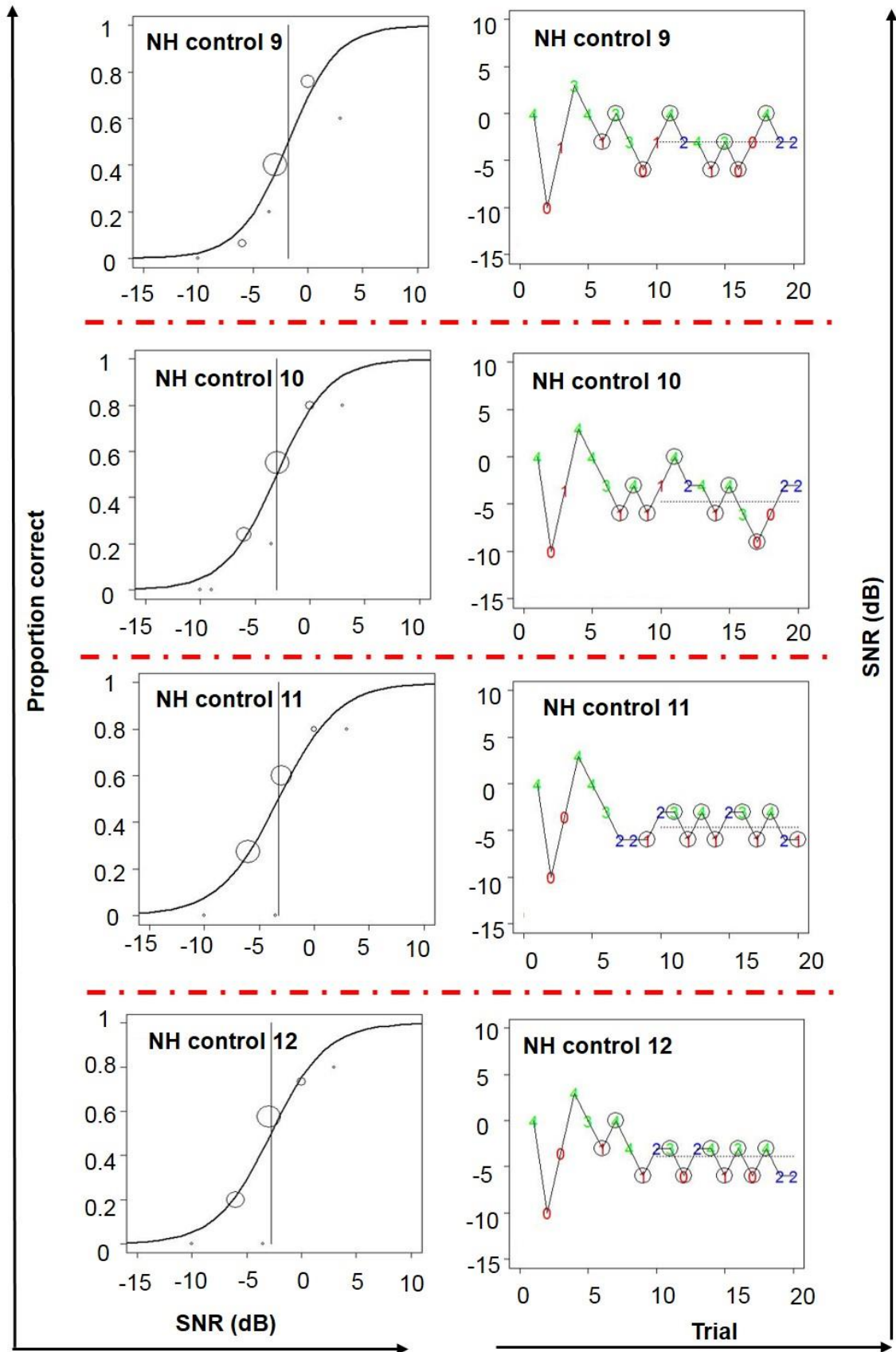
Therefore, as a result of these preliminary analyses, percentage of targets detected (with correction of over-reporting) was implemented as the chosen method for assessing visual accuracy in Experiments 1-8. However, monitoring of over-reporting was still executed, within a separate part of the finalised analyses, as part of the ongoing attempt to elucidate its potential role (if any) in relation to LE.



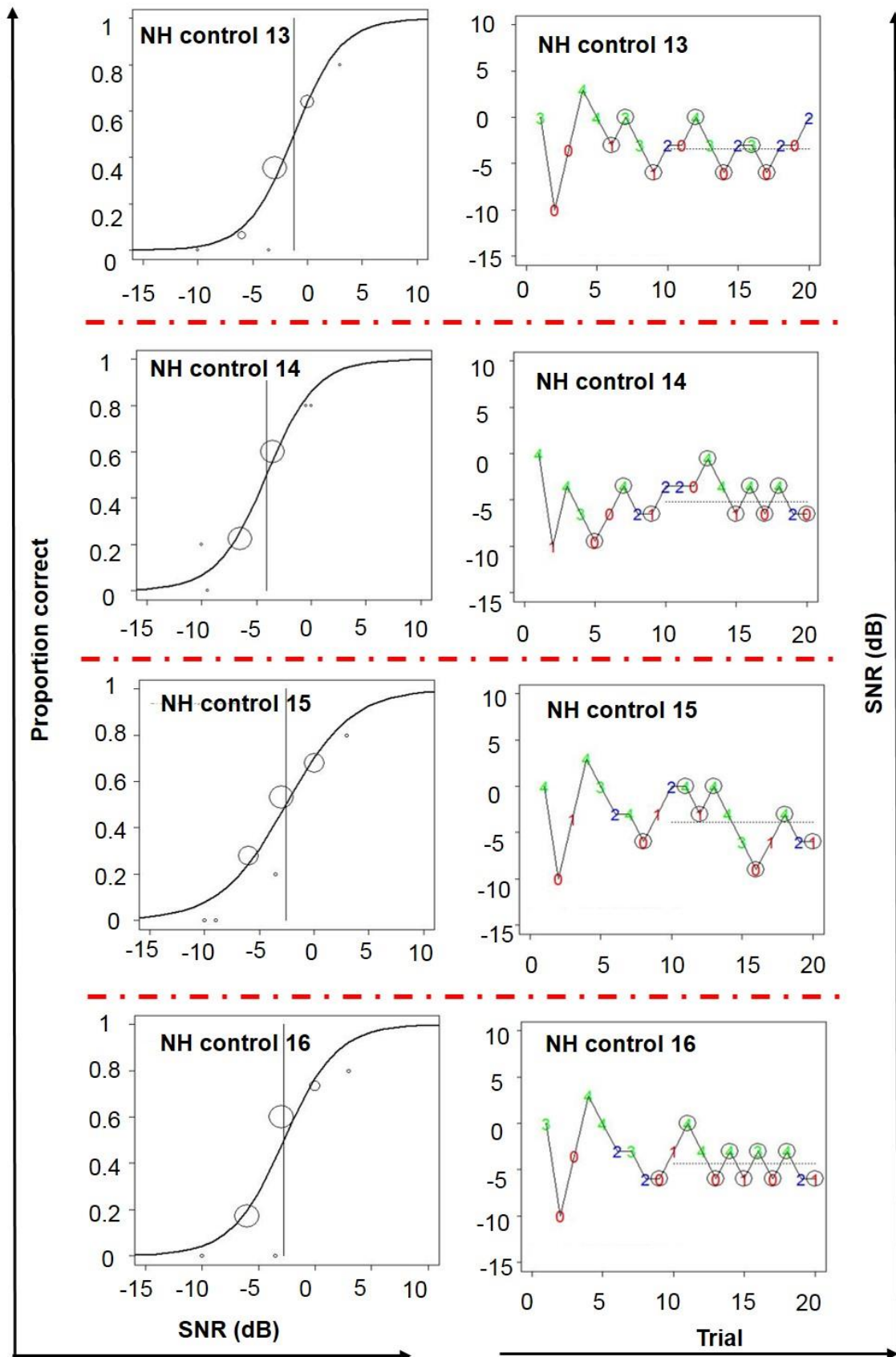
Appendix 4.1: The output from the adaptive noise tracking of NH controls 1-4 (Experiment 4). Left column displays all psychometric functions, with the proportion of average number of key words of a BEL sentence the participant successfully heard (and typed) at a given SNR. Superimposed circles on psychometric function = Participant's response data, with size of circle controlled by the number of trials run at that particular SNR. Right column displays all adaptive tracks, showing how the SNR was changed across the trials. Superimposed circles on adaptive track = Reversals. Superimposed numbers on adaptive track = Total number of key words correct, with colour indicating whether the track remained at the same SNR (blue), increased (red), or decreased (green) on the next trial.



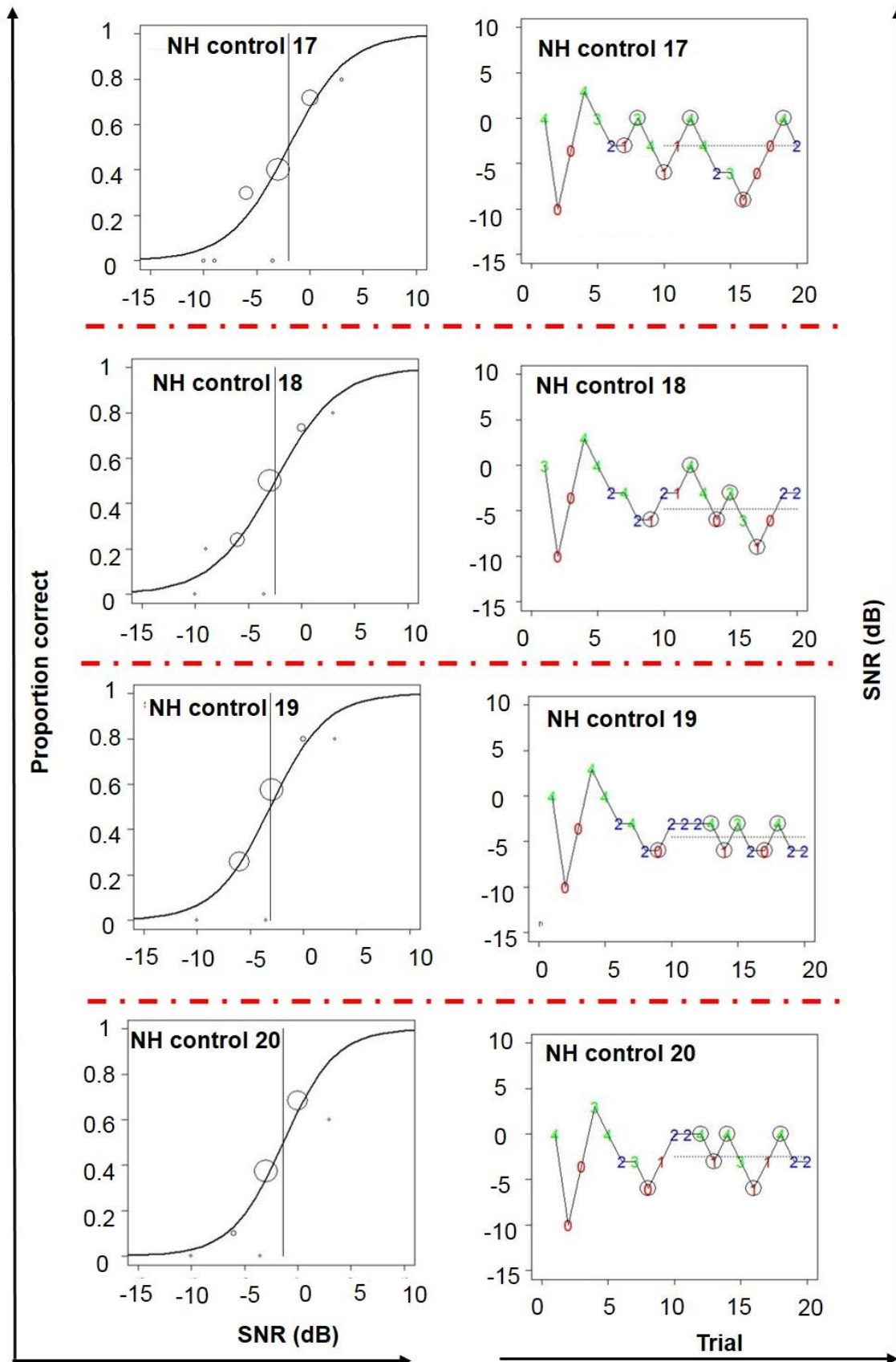
Appendix 4.2: As for Appendix 4.1, but output this time from NH controls 5-8 (Experiment 4).

Appendix 4.3: As for Appendix 4.1, but output this time from NH controls 9-12 (Experiment 4).

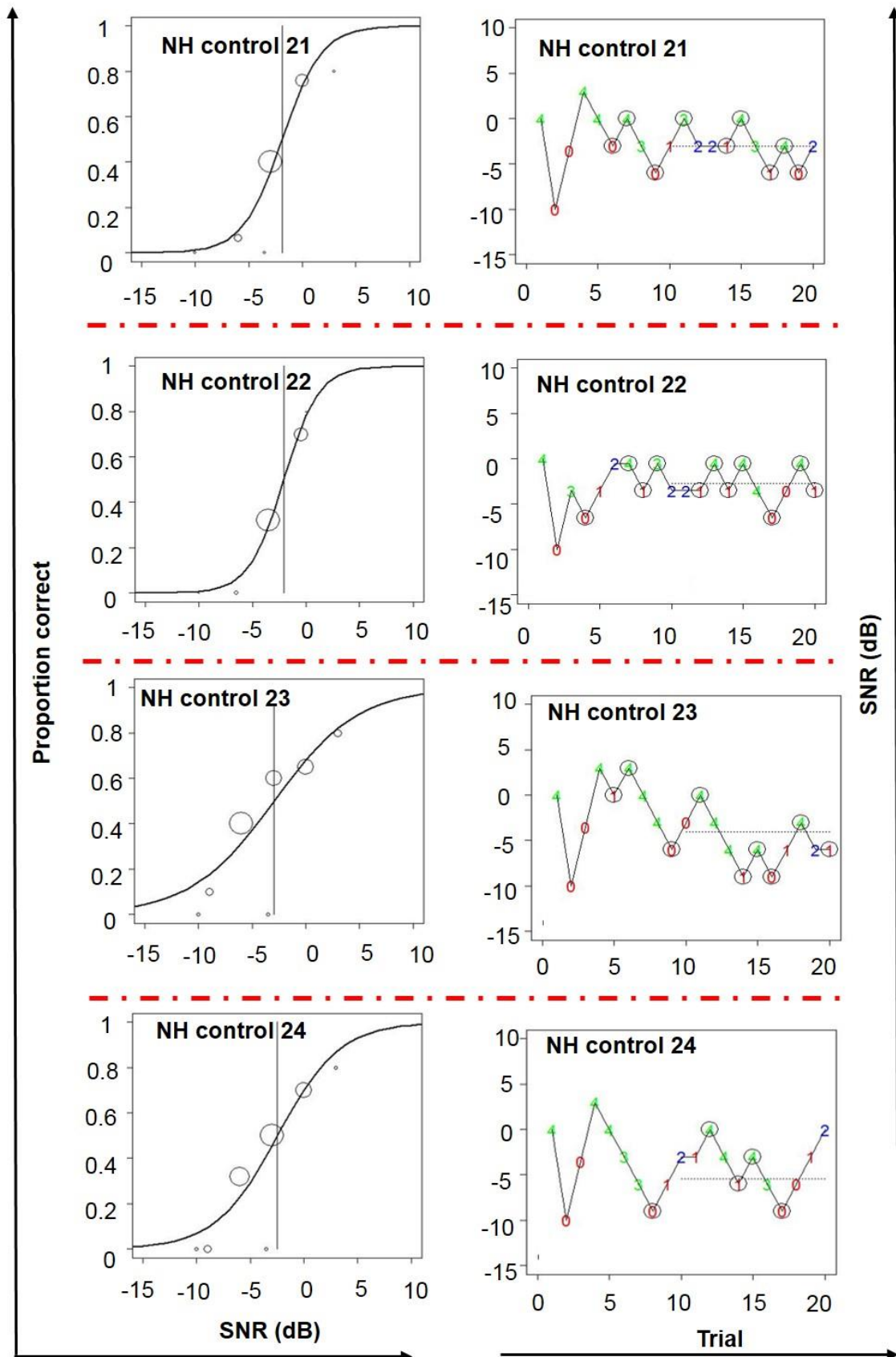
Appendix 4.4: As for Appendix 4.1, but output this time from NH controls 13-16 (Experiment 4).



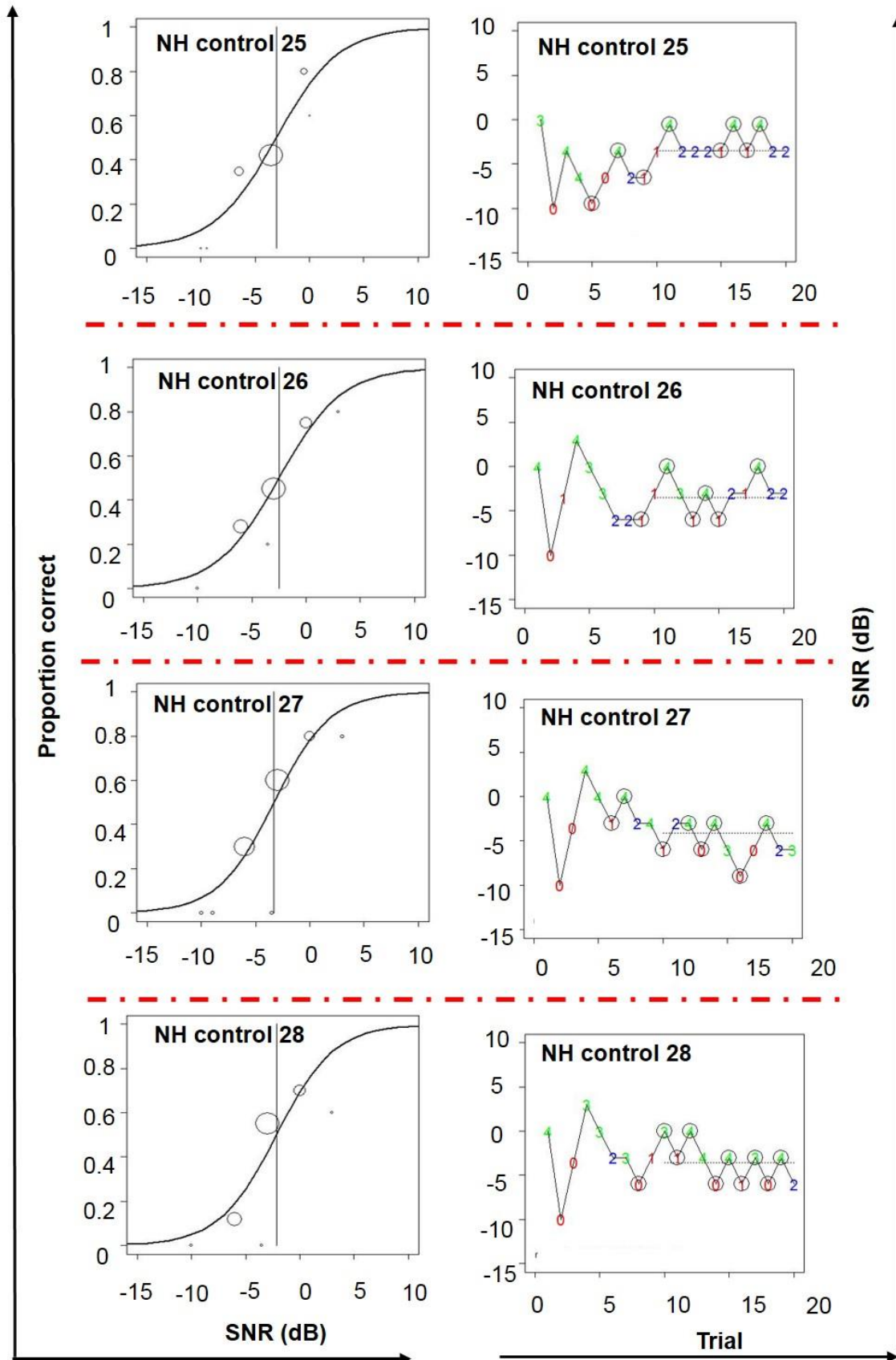
Appendix 4.5: As for Appendix 4.1, but output this time from NH controls 17-20 (Experiment 4).



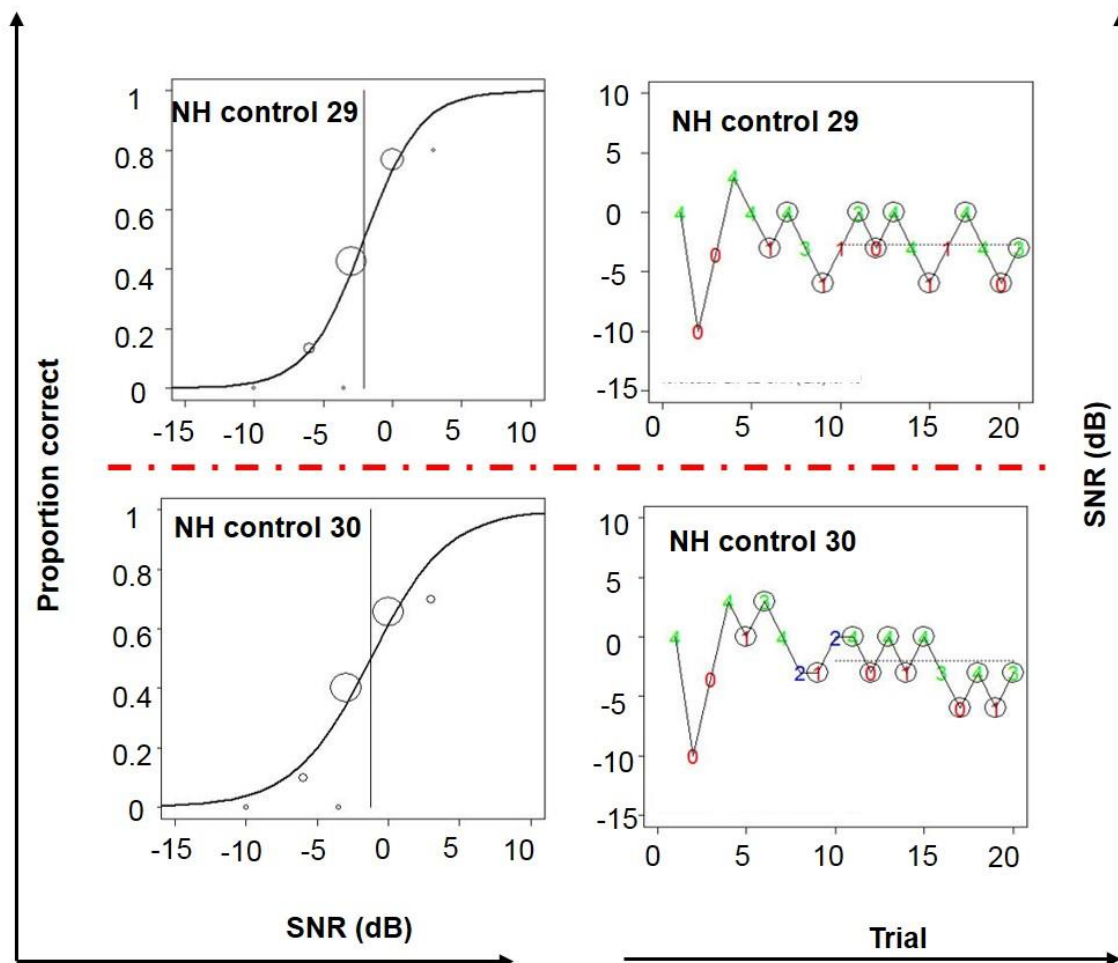
Appendix 4.6: As for Appendix 4.1, but output this time from NH controls 21-24 (Experiment 4).



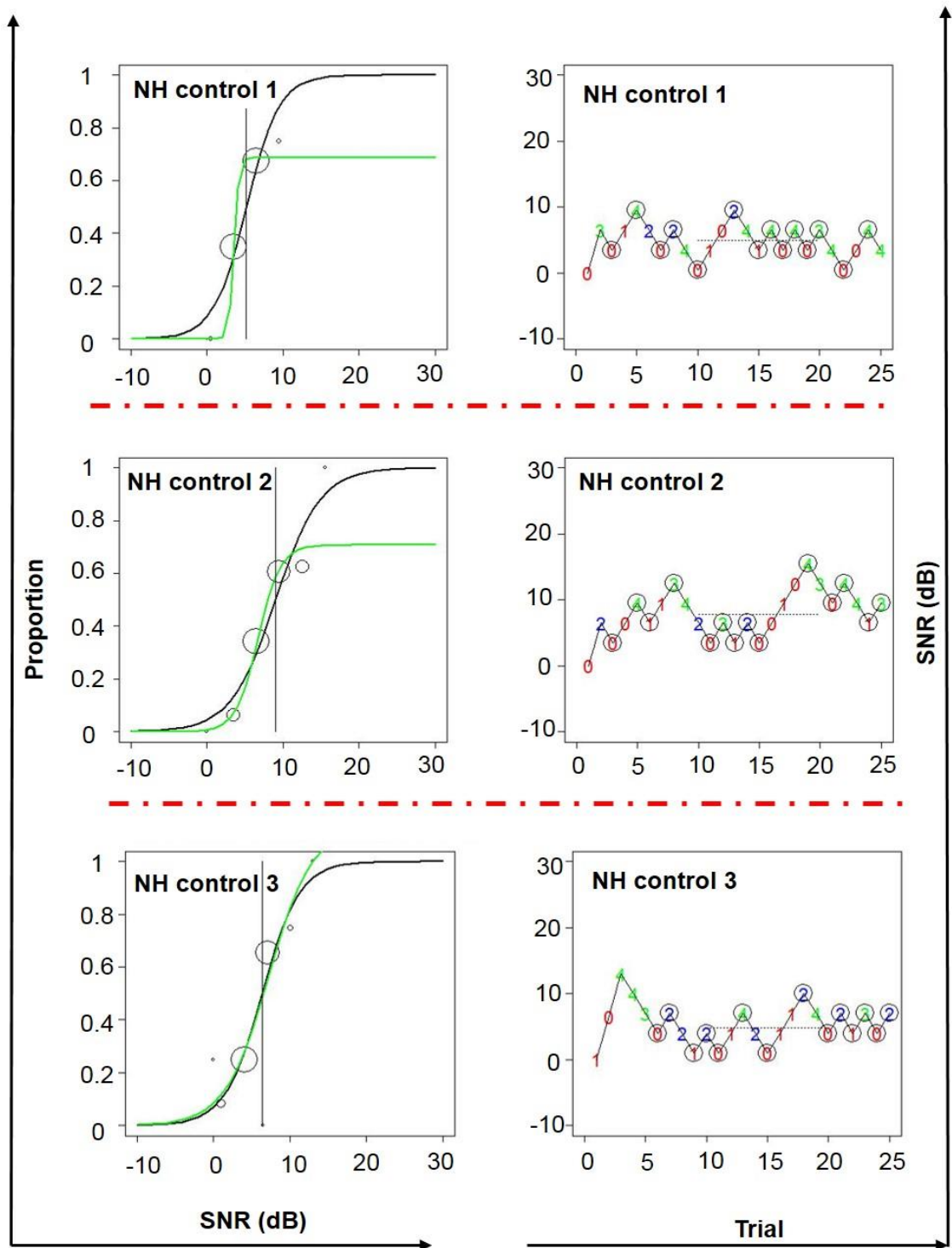
Appendix 4.7: As for Appendix 4.1, but output this time from NH controls 25-28 (Experiment 4).

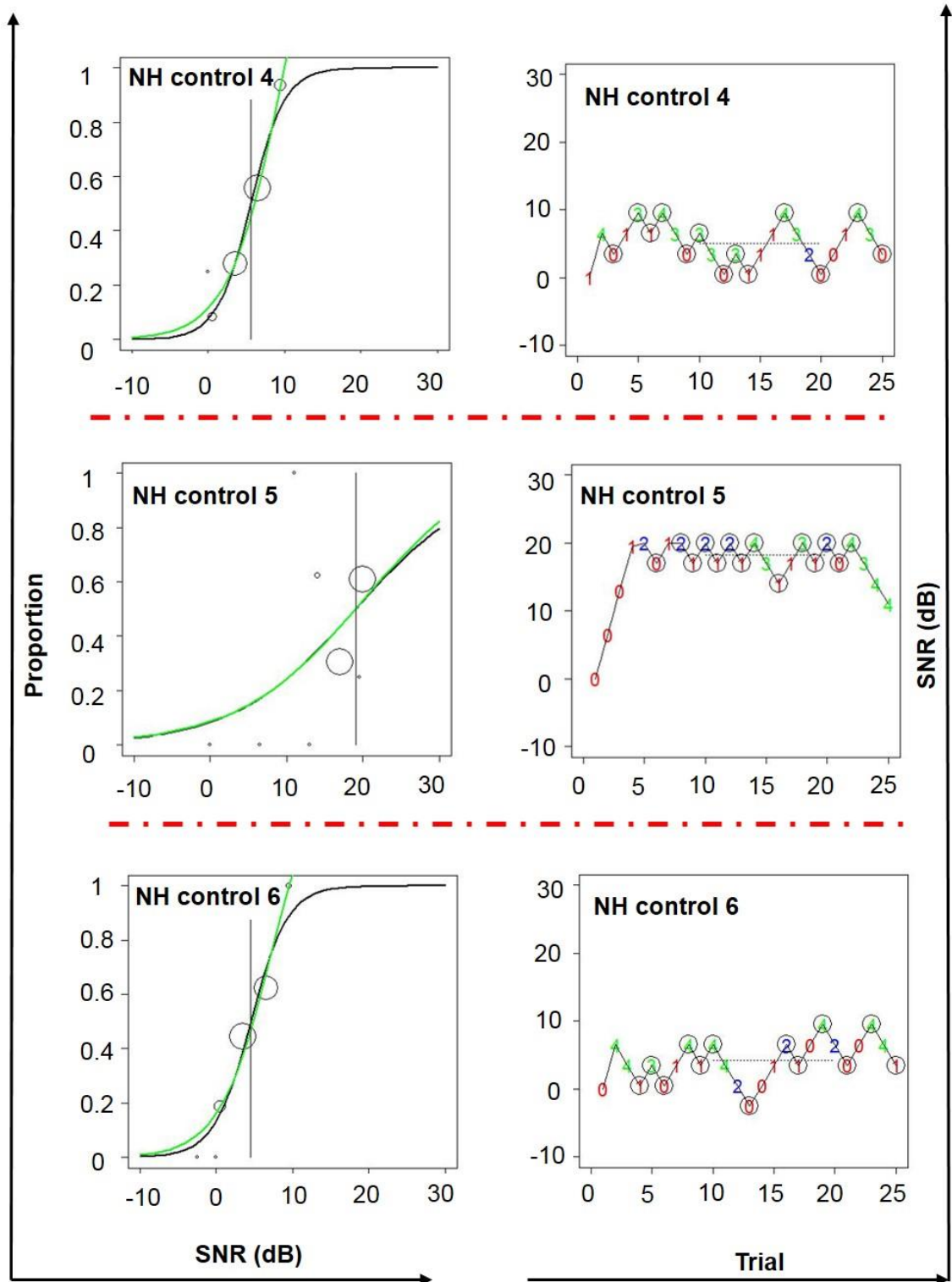


Appendix 4.8: As for Appendix 4.1, but output this time from NH controls 29 and 30 (Experiment 4).

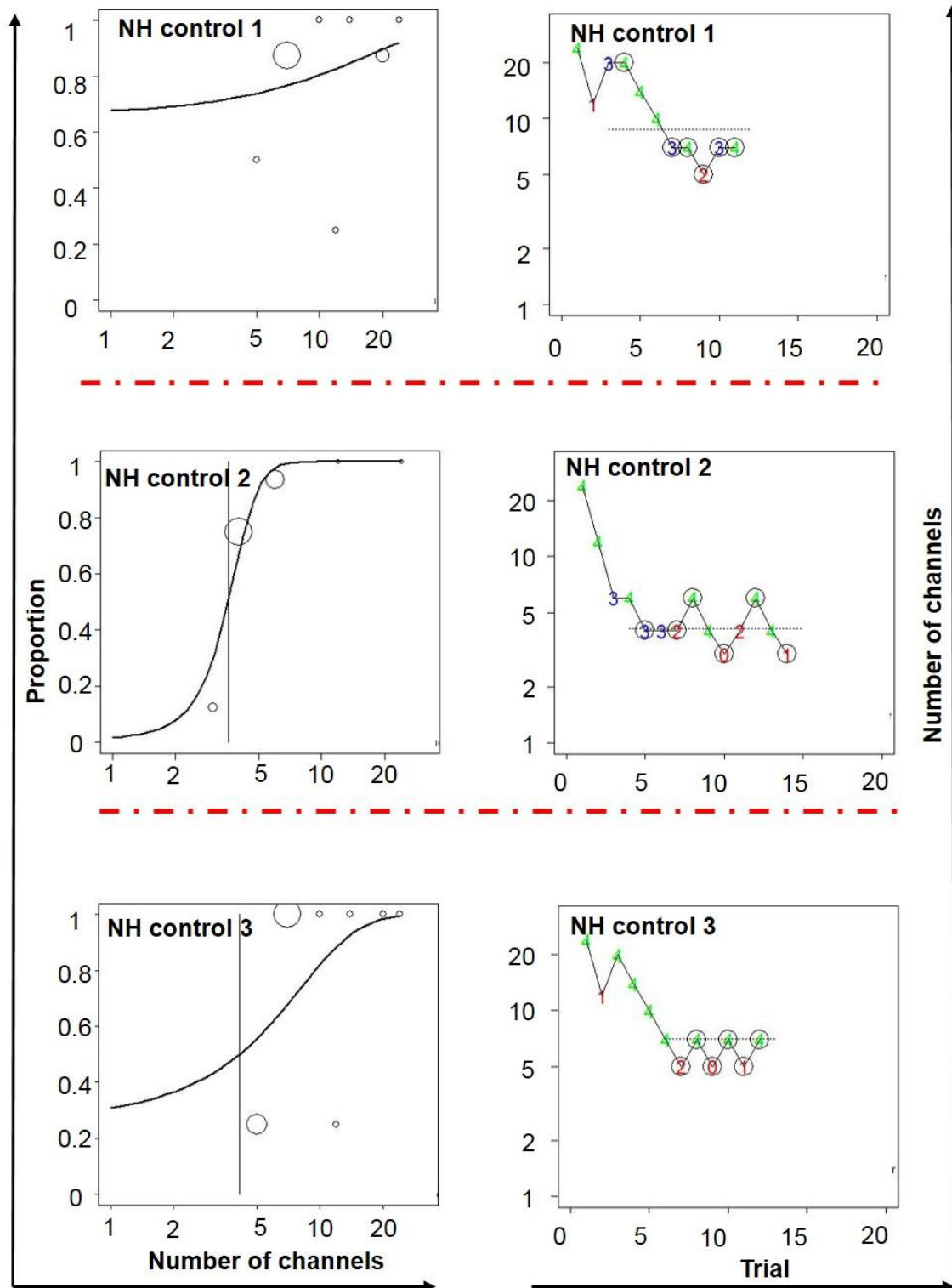


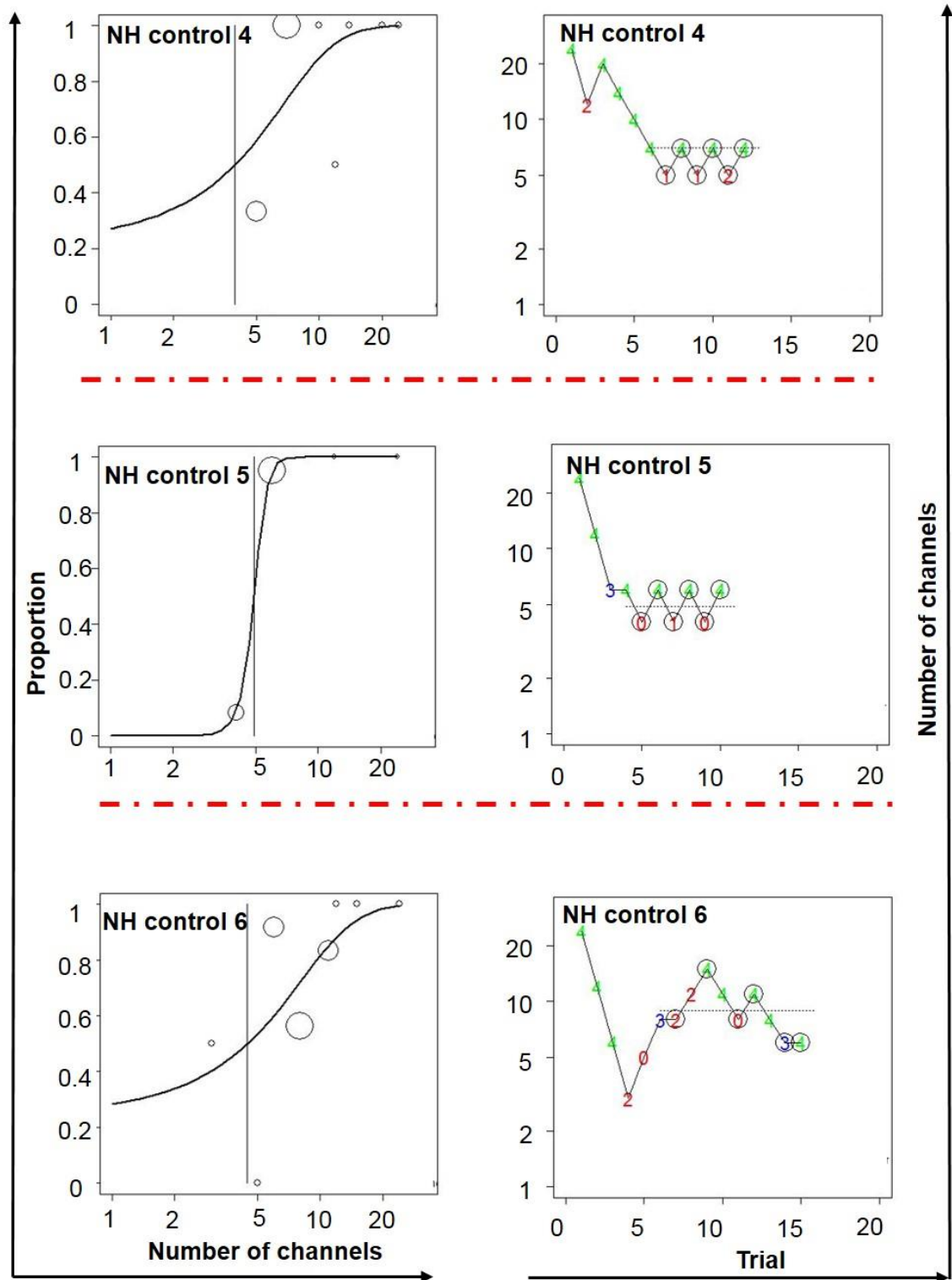
Appendix 5.1: As for Appendix 4.1, but output this time from Experiment 5 for NH controls 1-3. There is the addition of a green line, which indicates the logistic regression with additional parameter in the fit, i.e. an upper asymptote, otherwise known as lapse rate (to enable psychometric function to plateau at $y < 1$).



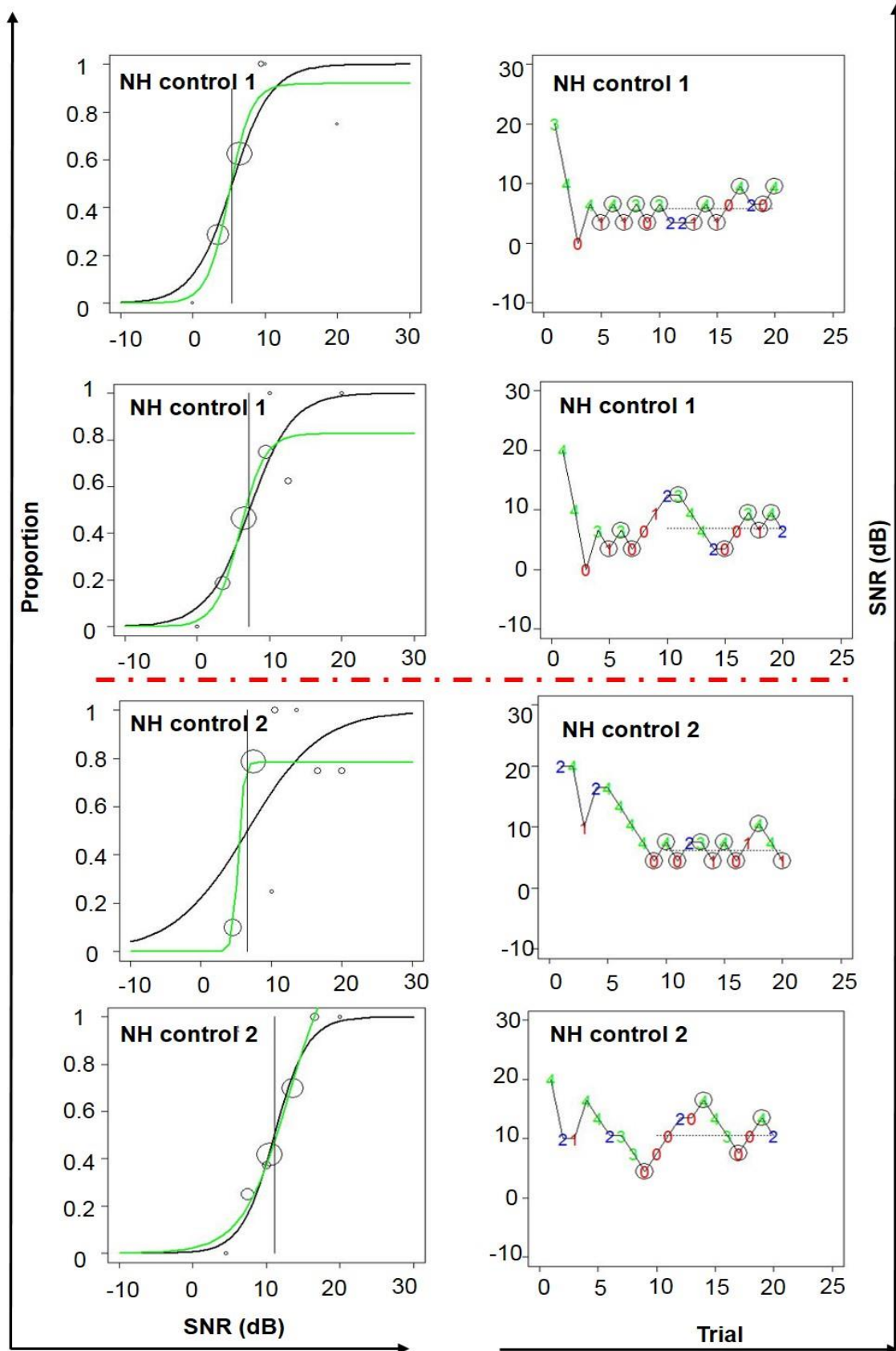
Appendix 5.2: As for Appendix 5.1, but output this time from NH controls 4-6 (Experiment 5).

Appendix 5.3: As for Appendix 5.1, but output this time from the adaptive channel tracking of NH controls 1-3 (Experiment 5). Psychometric functions and adaptive tracks relate to performance at a given channel number instead of SNR. Green line is not applicable.

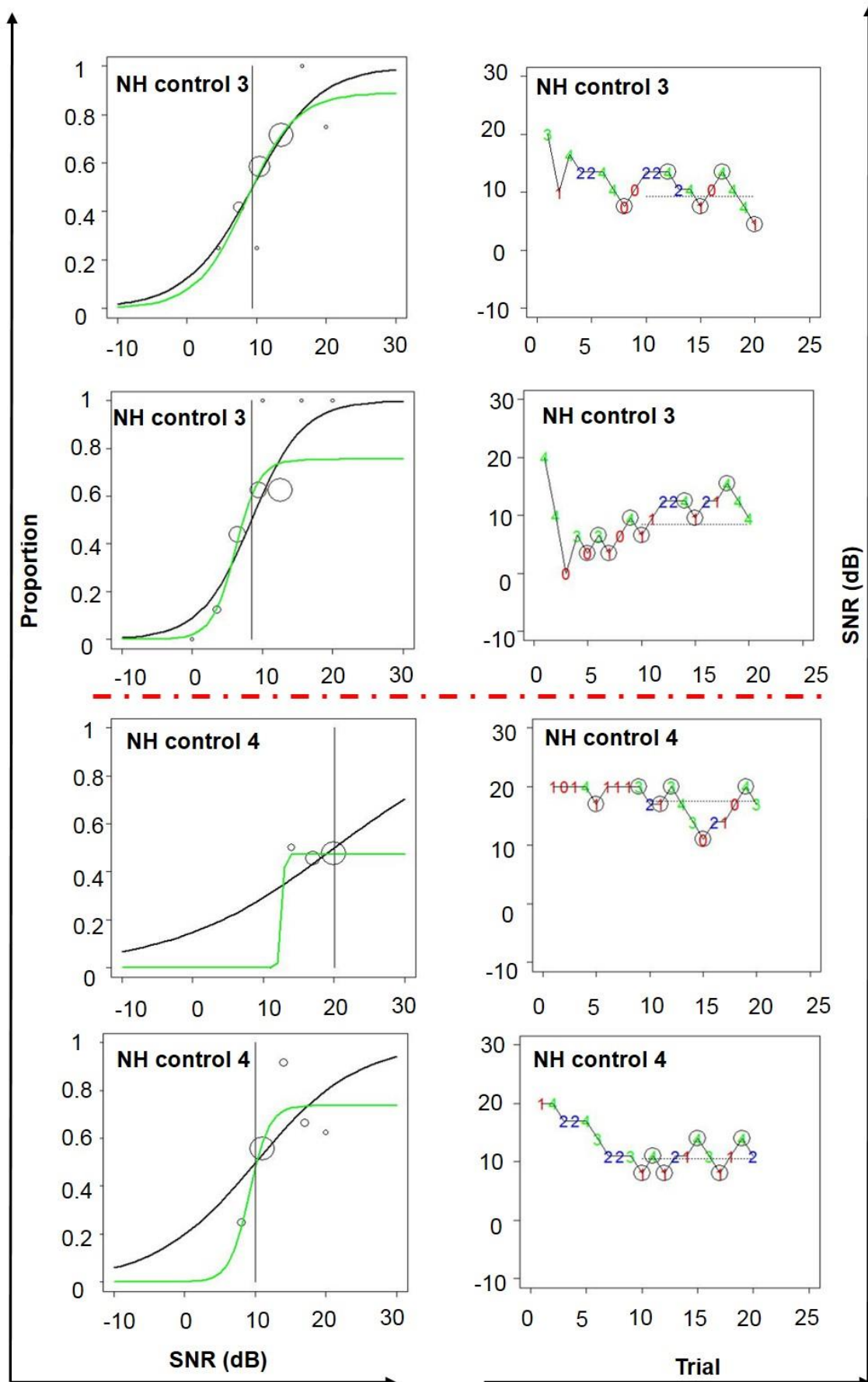


Appendix 5.4: As for Appendix 5.3, but output this time from NH controls 4-6 (Experiment 5).

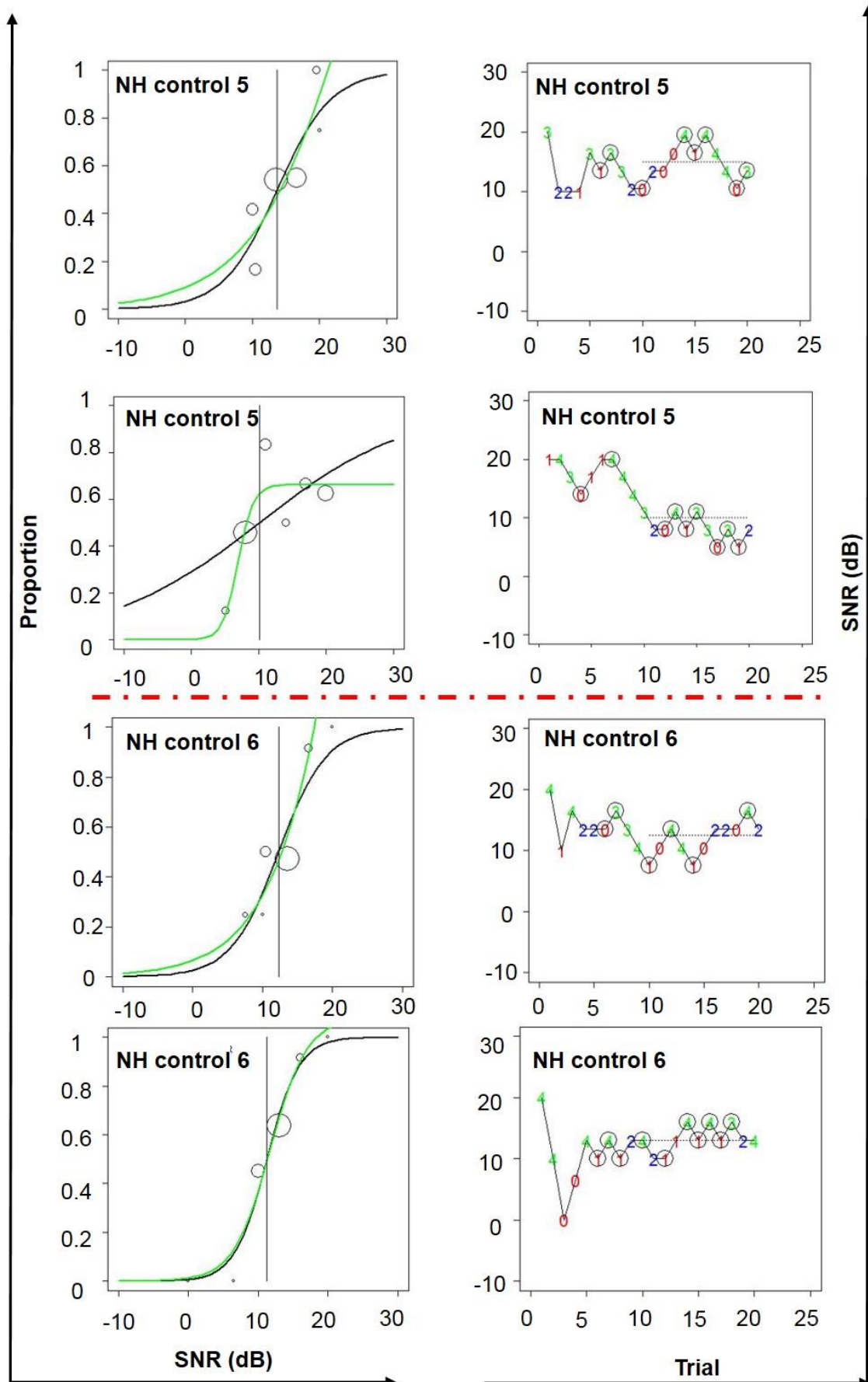
Appendix 5.5: As for Appendix 5.1, but output this time from Experiment 6 for NH controls 1 & 2 (in both attempts). Green line is now applicable.



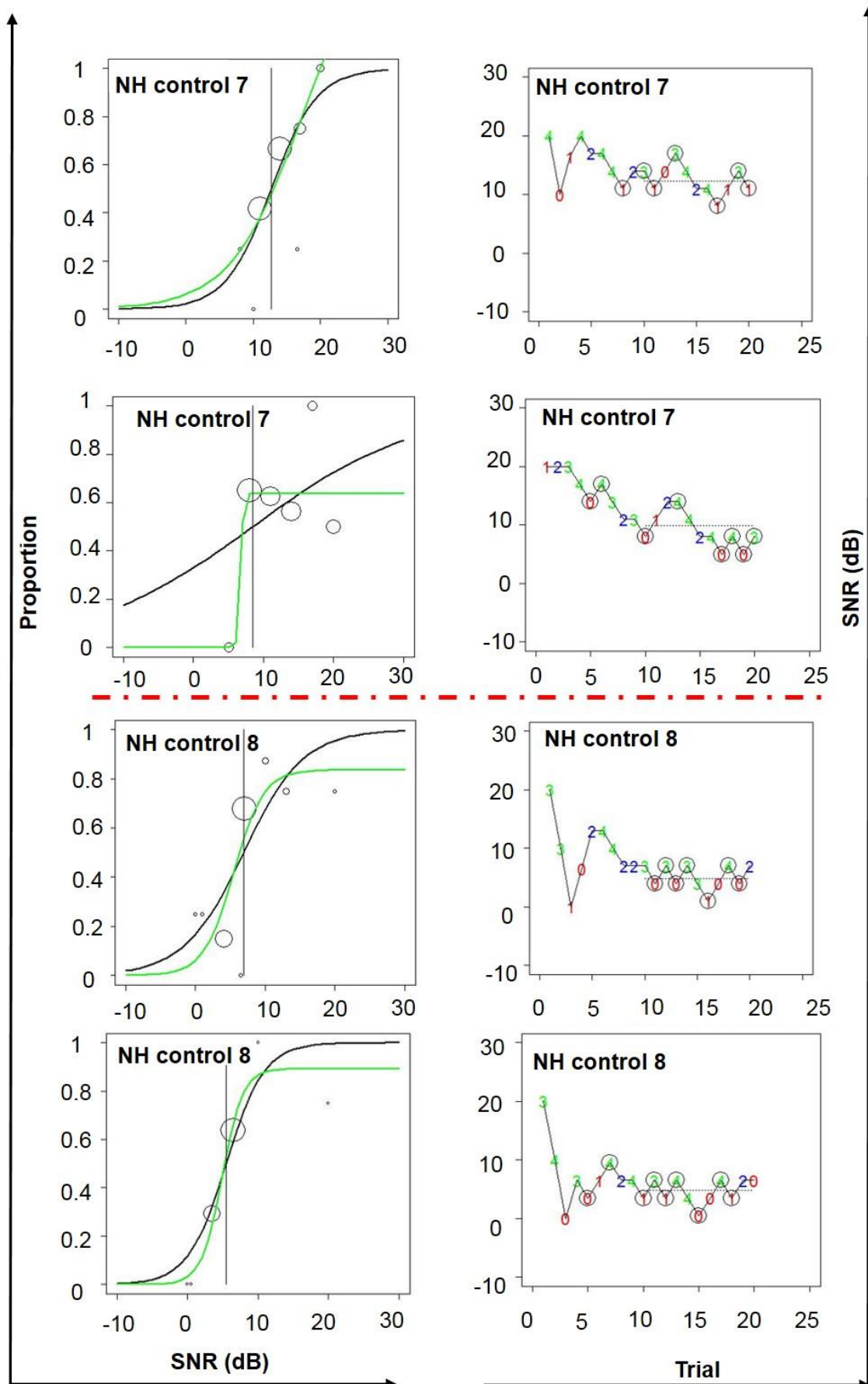
Appendix 5.6: As for Appendix 5.5, but output this time from NH controls 3 & 4 (Experiment 6).



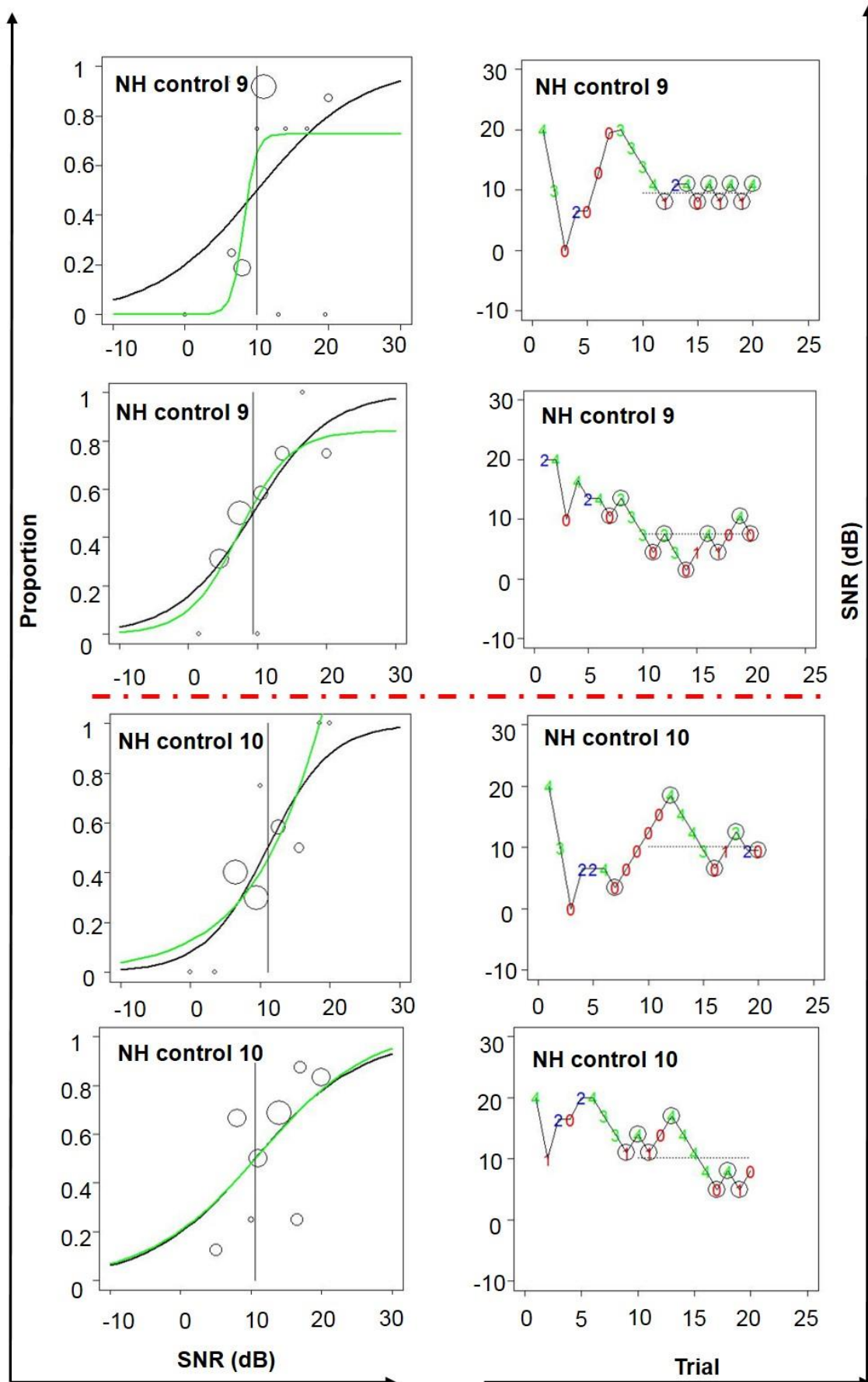
Appendix 5.7: As for Appendix 5.5, but output this time from NH controls 5 & 6 (Experiment 6).



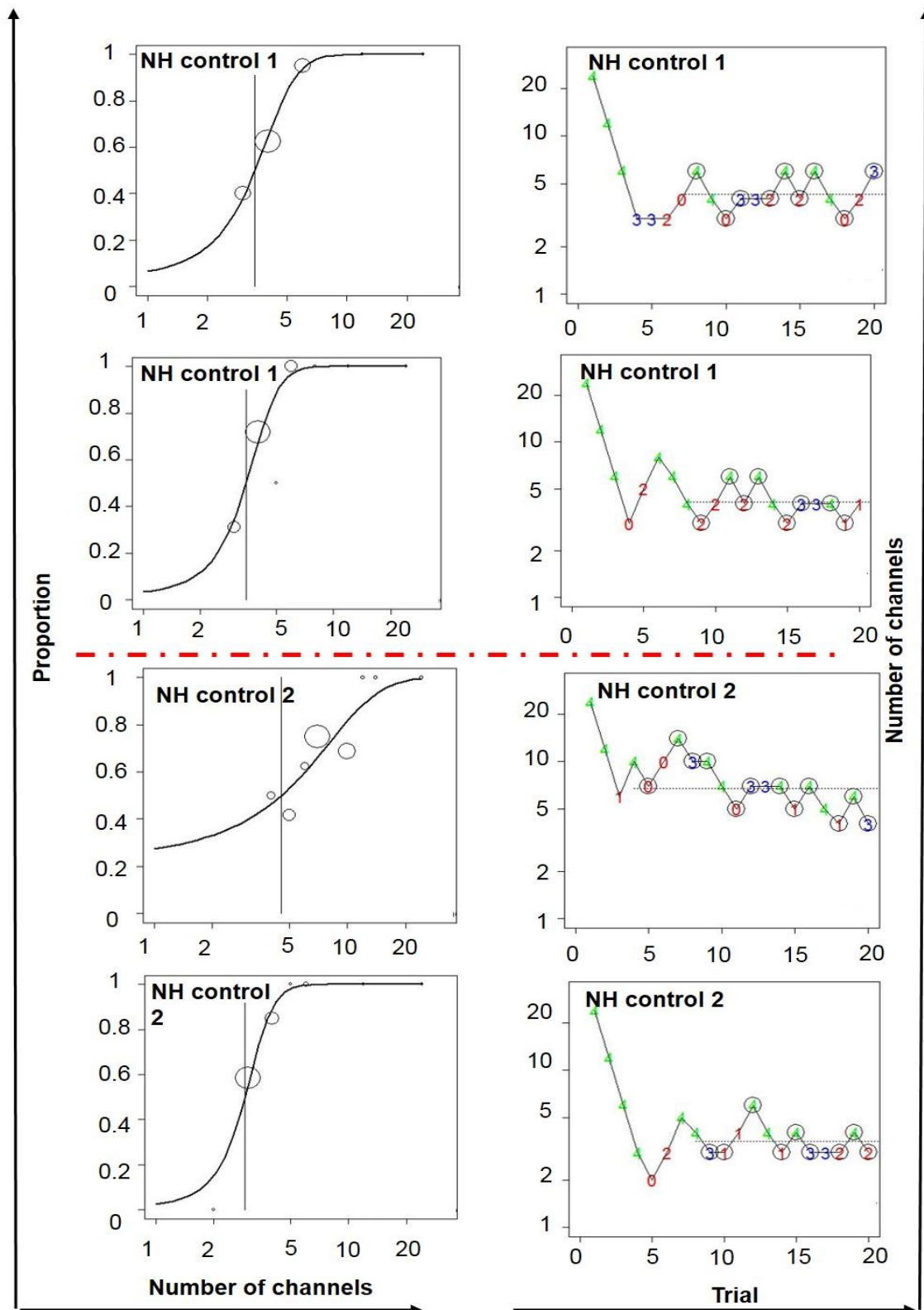
Appendix 5.8: As for Appendix 5.5, but output this time from NH controls 7 & 8 (Experiment 6).



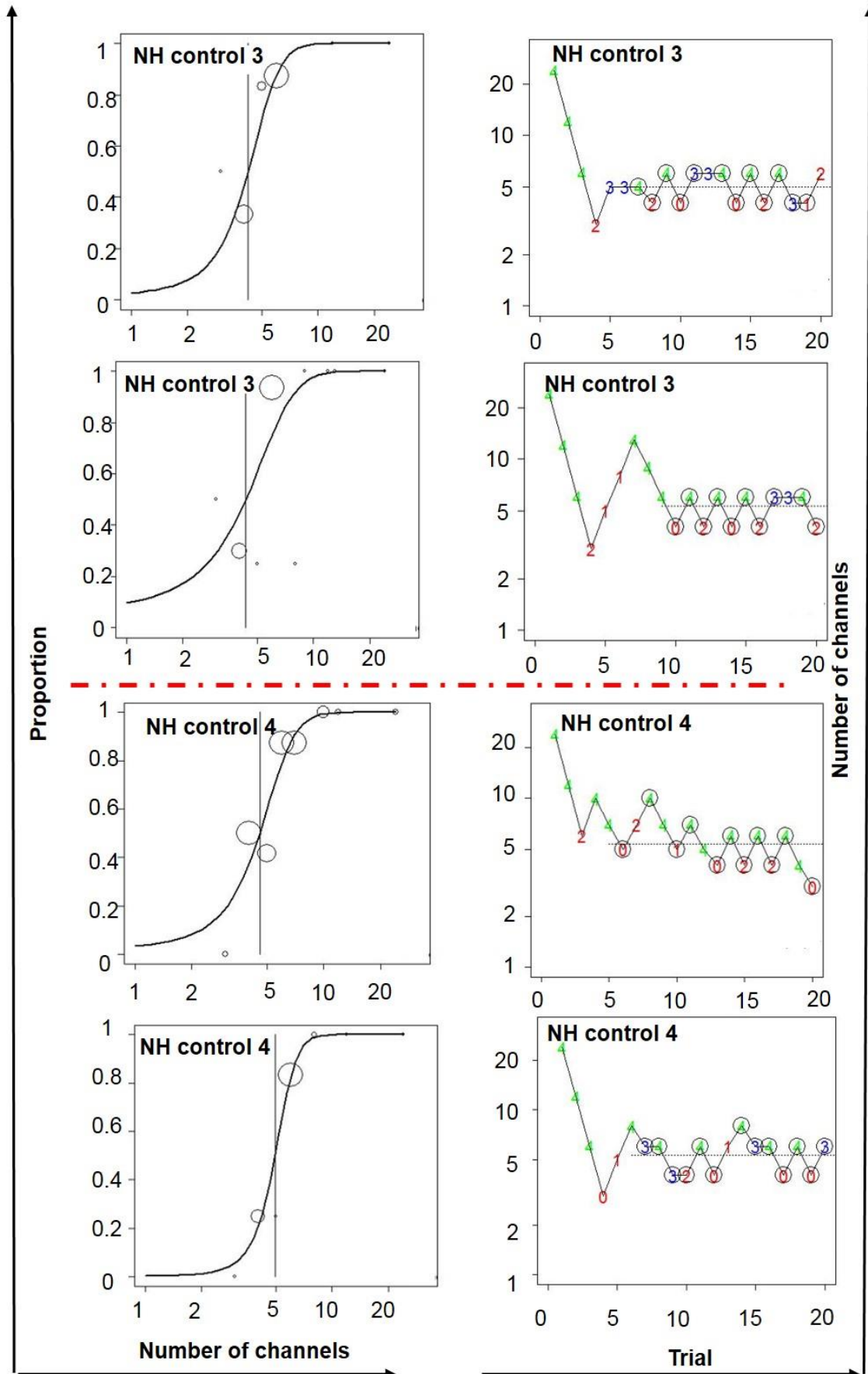
Appendix 5.9: As for Appendix 5.5, but output this time from NH controls 9 & 10 (Experiment 6).



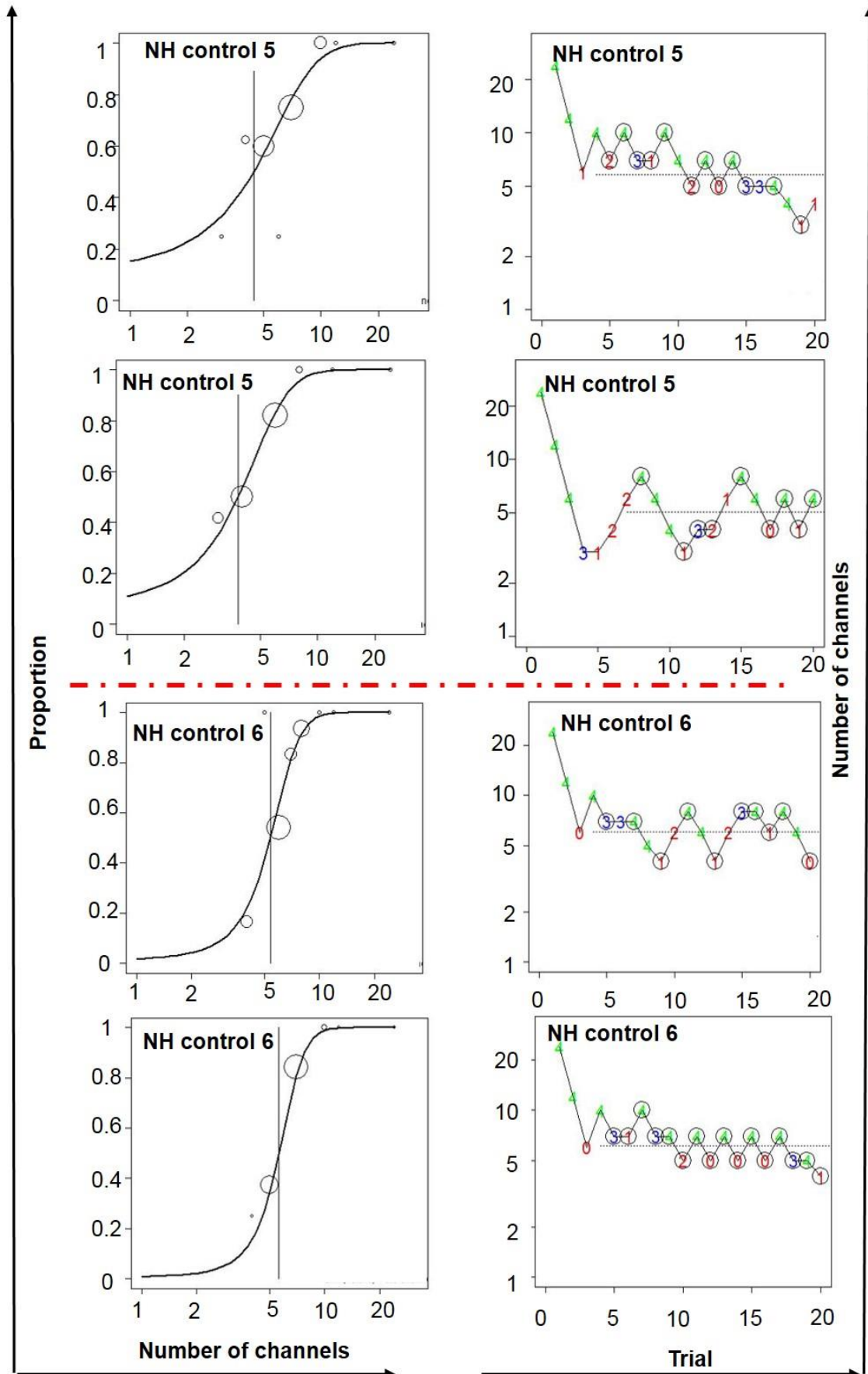
Appendix 5.10: As for Appendix 5.3, but output this time from Experiment 6 for NH controls 1 & 2 (in both attempts). Green line is not applicable.



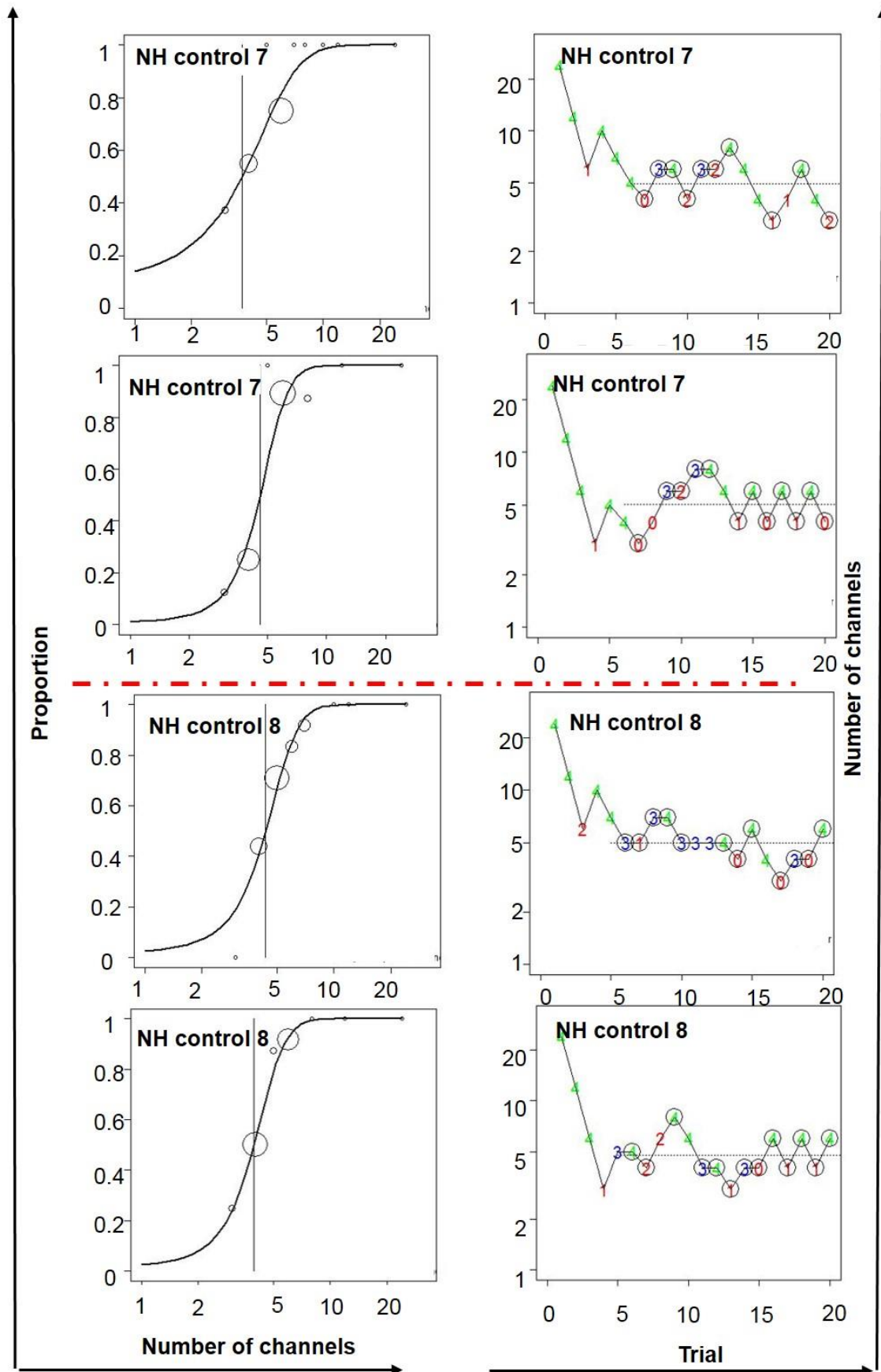
Appendix 5.11: As for Appendix 5.10, but output this time from NH controls 3 & 4 (Experiment 6).



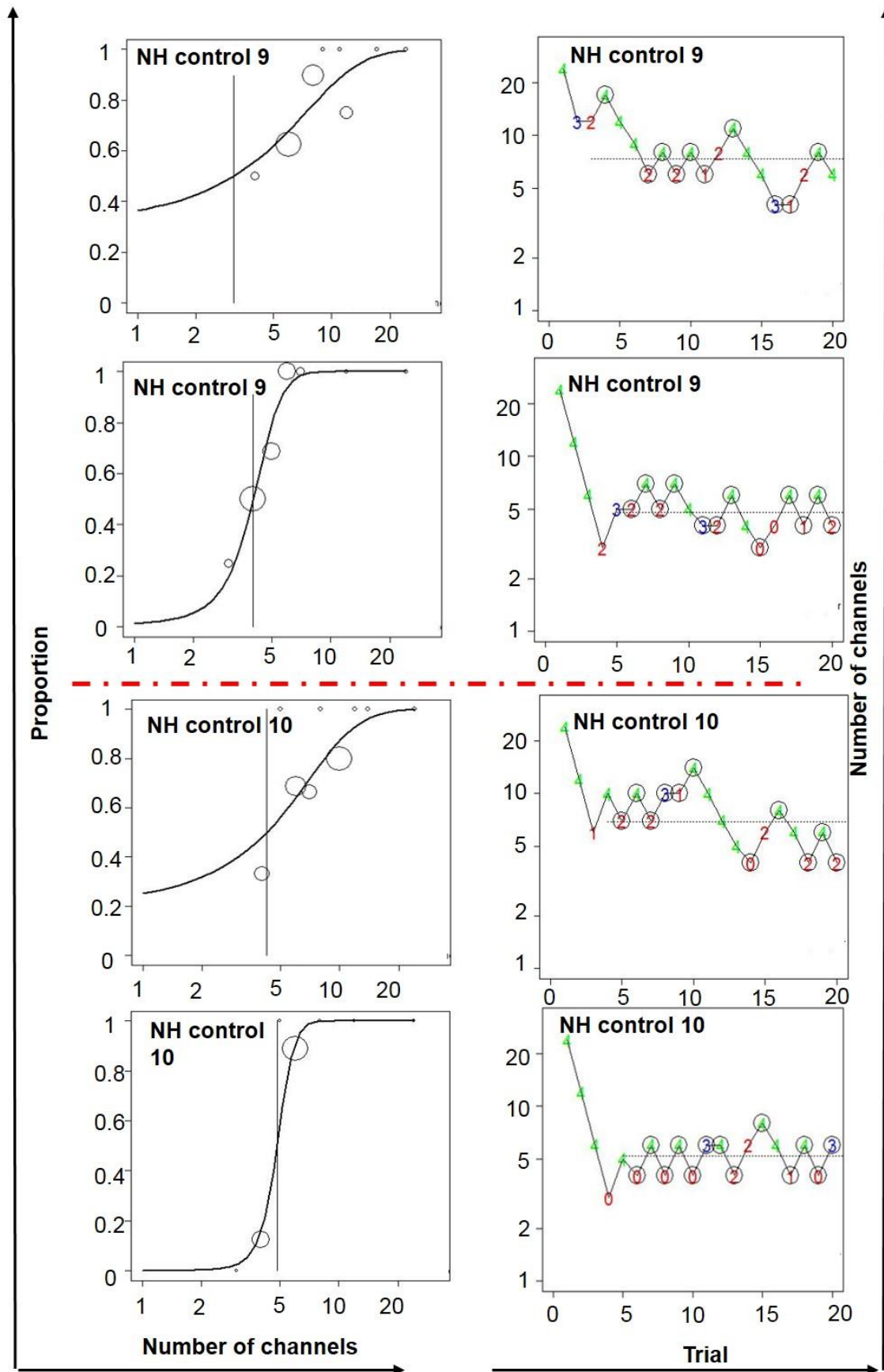
Appendix 5.12: As for Appendix 5.10, but output this time from NH controls 5 & 6 (Experiment 6).



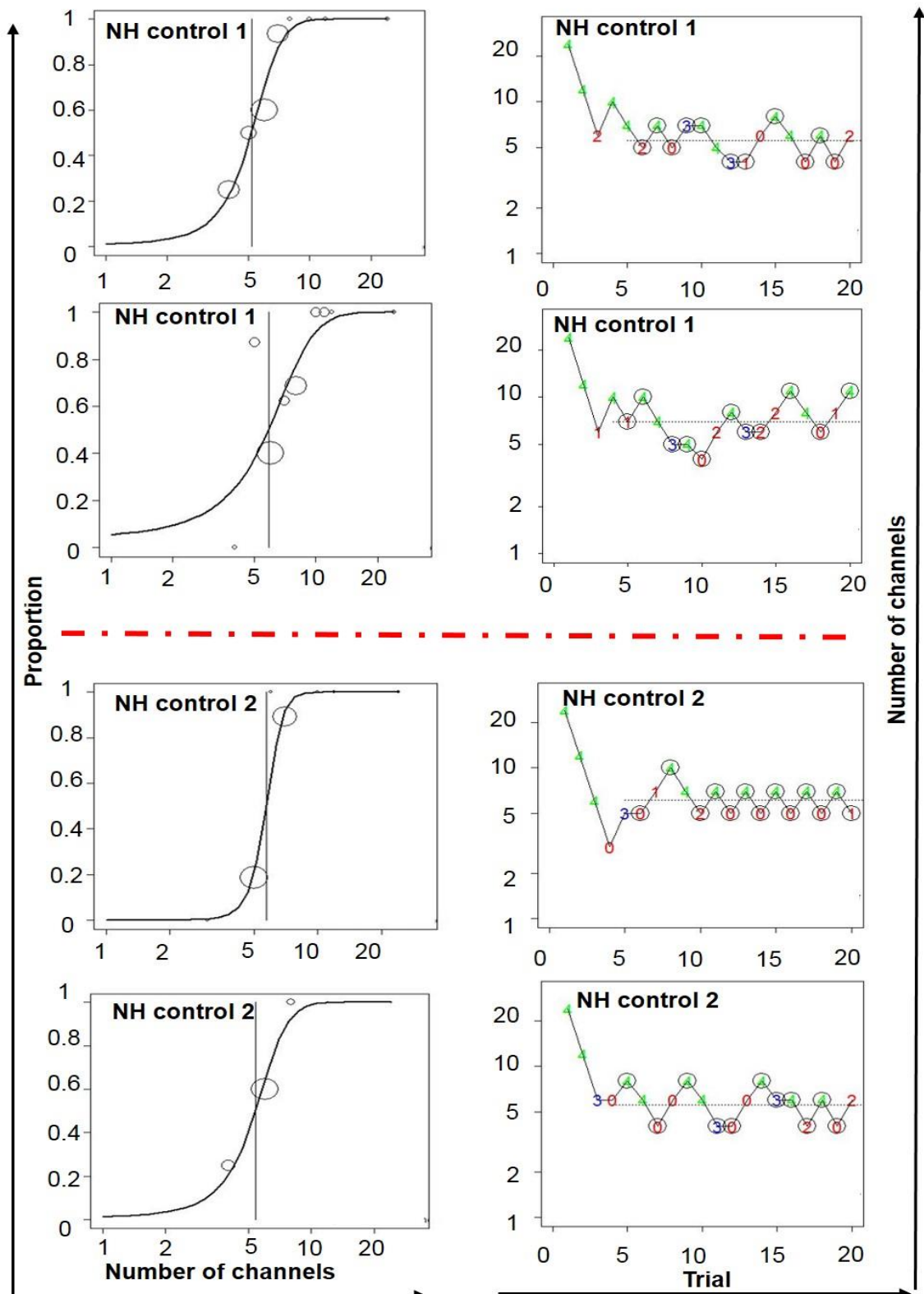
Appendix 5.13: As for Appendix 5.10, but output this time from NH controls 7 & 8 (Experiment 6).



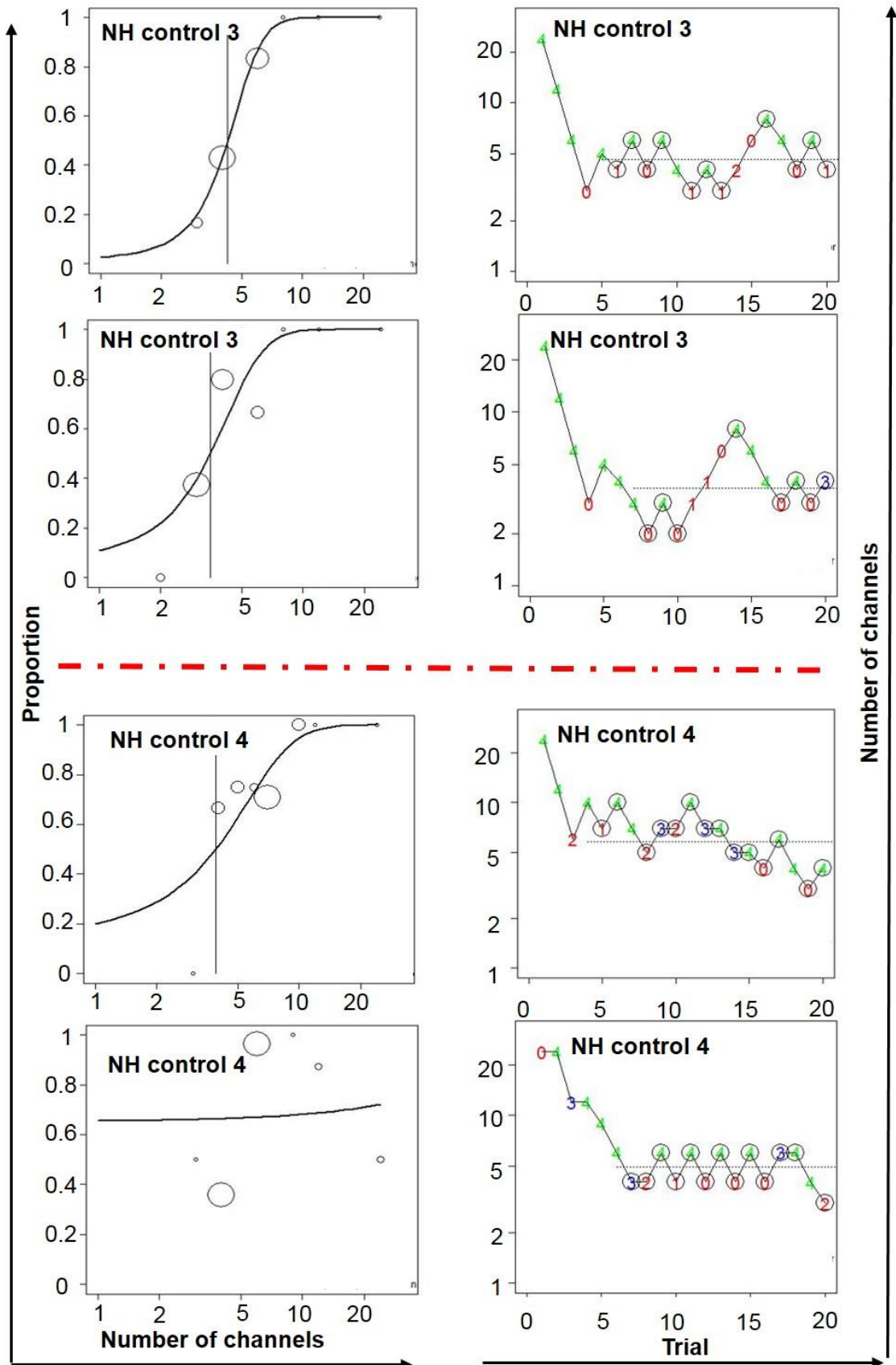
Appendix 5.14: As for Appendix 5.10, but output this time from NH controls 9 & 10 (Experiment 6).



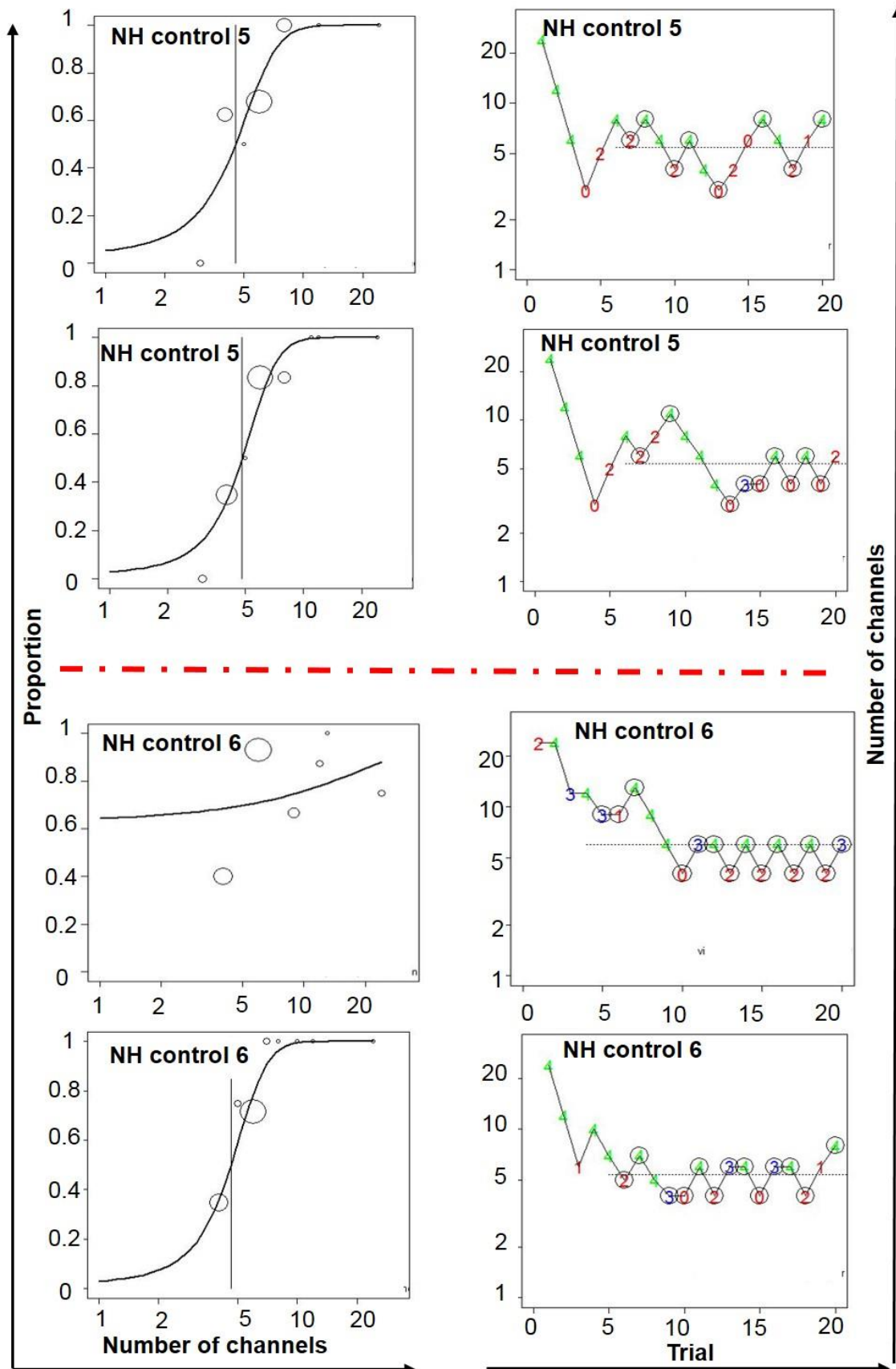
Appendix 5.15: As for Appendix 5.3, but output this time from Experiment 7 for NH controls 1 & 2 (in both attempts). Green line is not applicable.



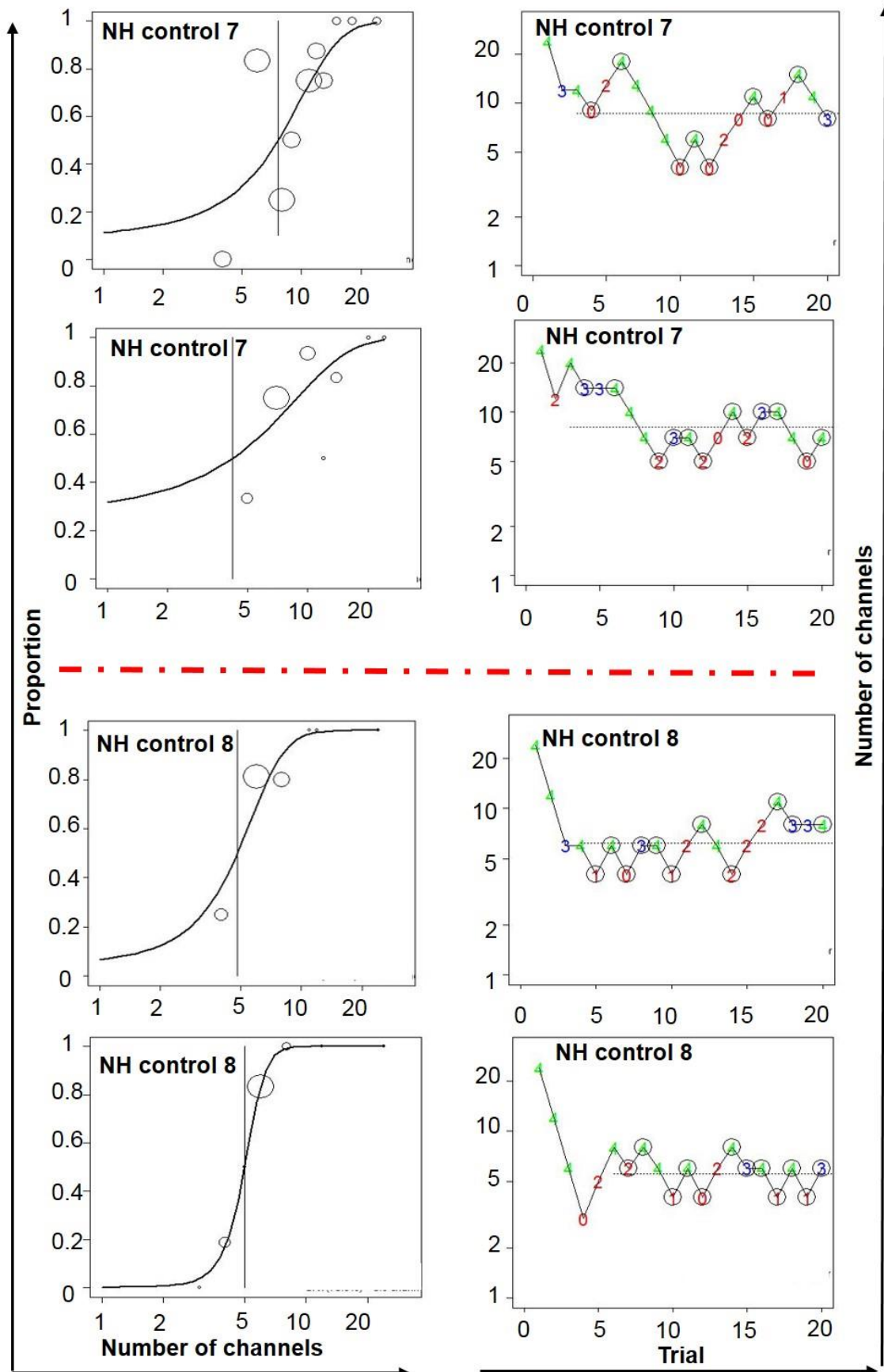
Appendix 5.16: As for Appendix 5.15, but output this time from NH controls 3 & 4 (Experiment 7).



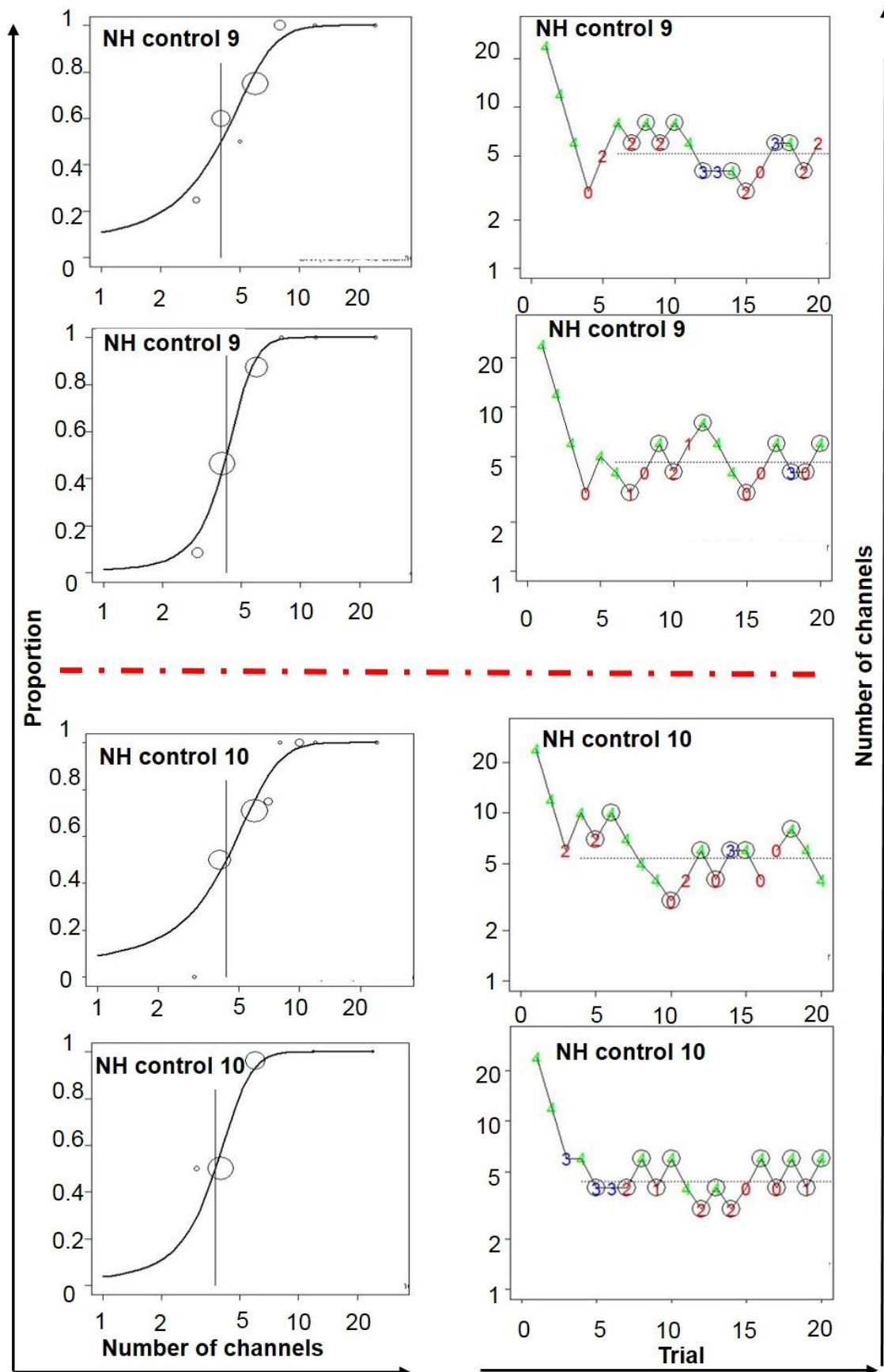
Appendix 5.17: As for Appendix 5.15, but output this time from NH controls 5 & 6 (Experiment 7).



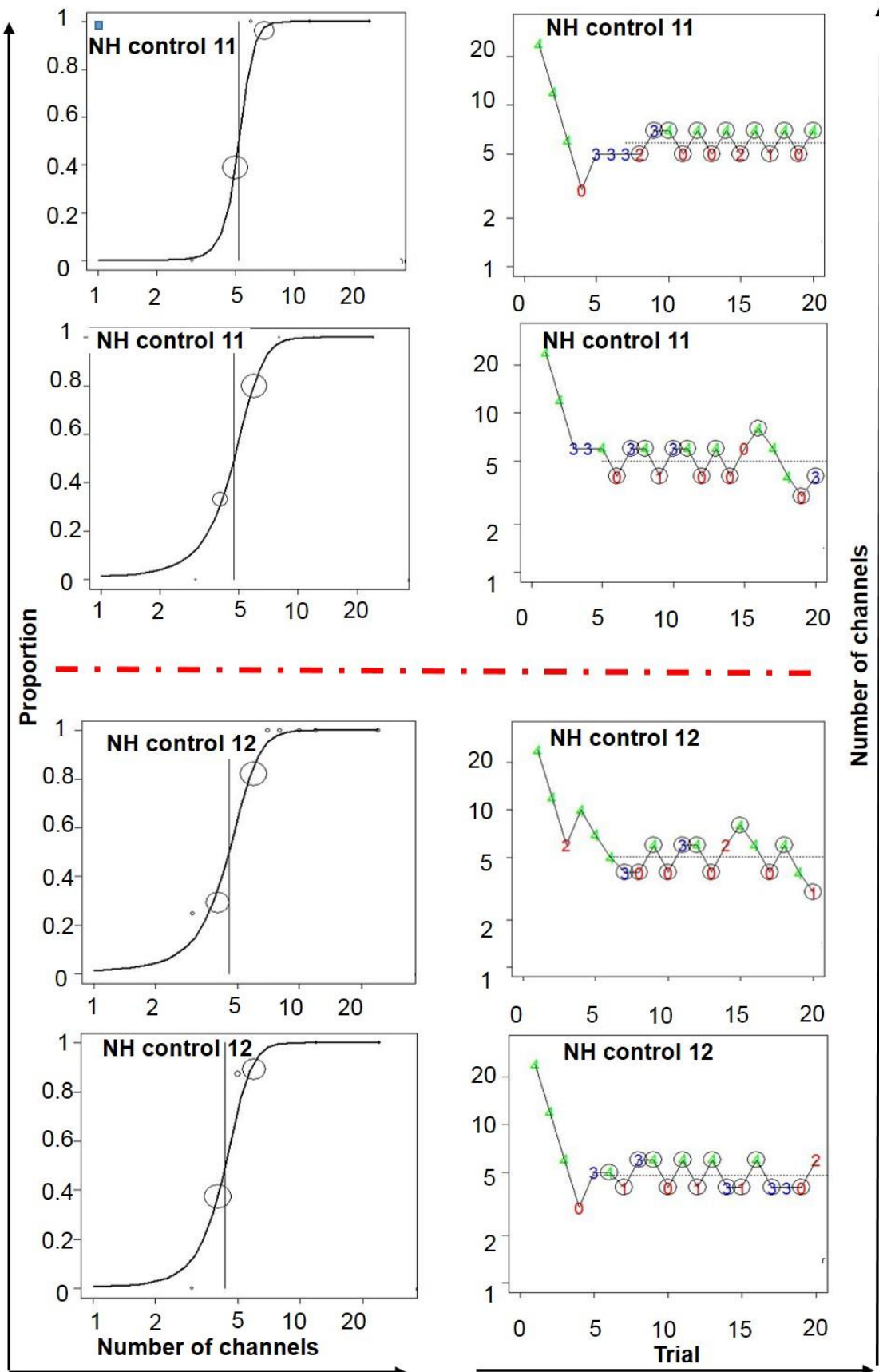
Appendix 5.18: As for Appendix 5.15, but output this time from NH controls 7 & 8 (Experiment 7).



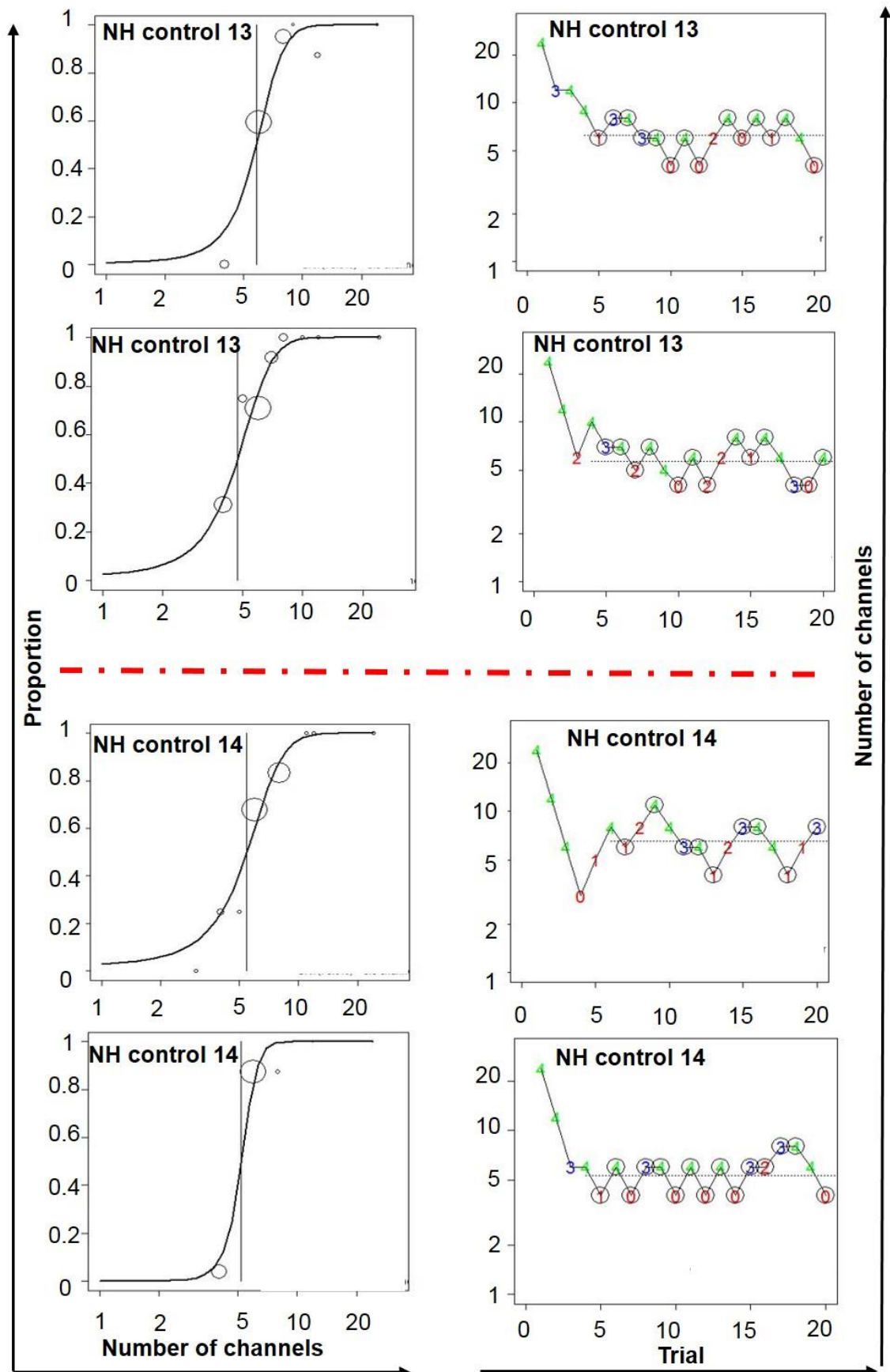
Appendix 5.19: As for Appendix 5.15, but output this time from NH controls 9 & 10 (Experiment 7).



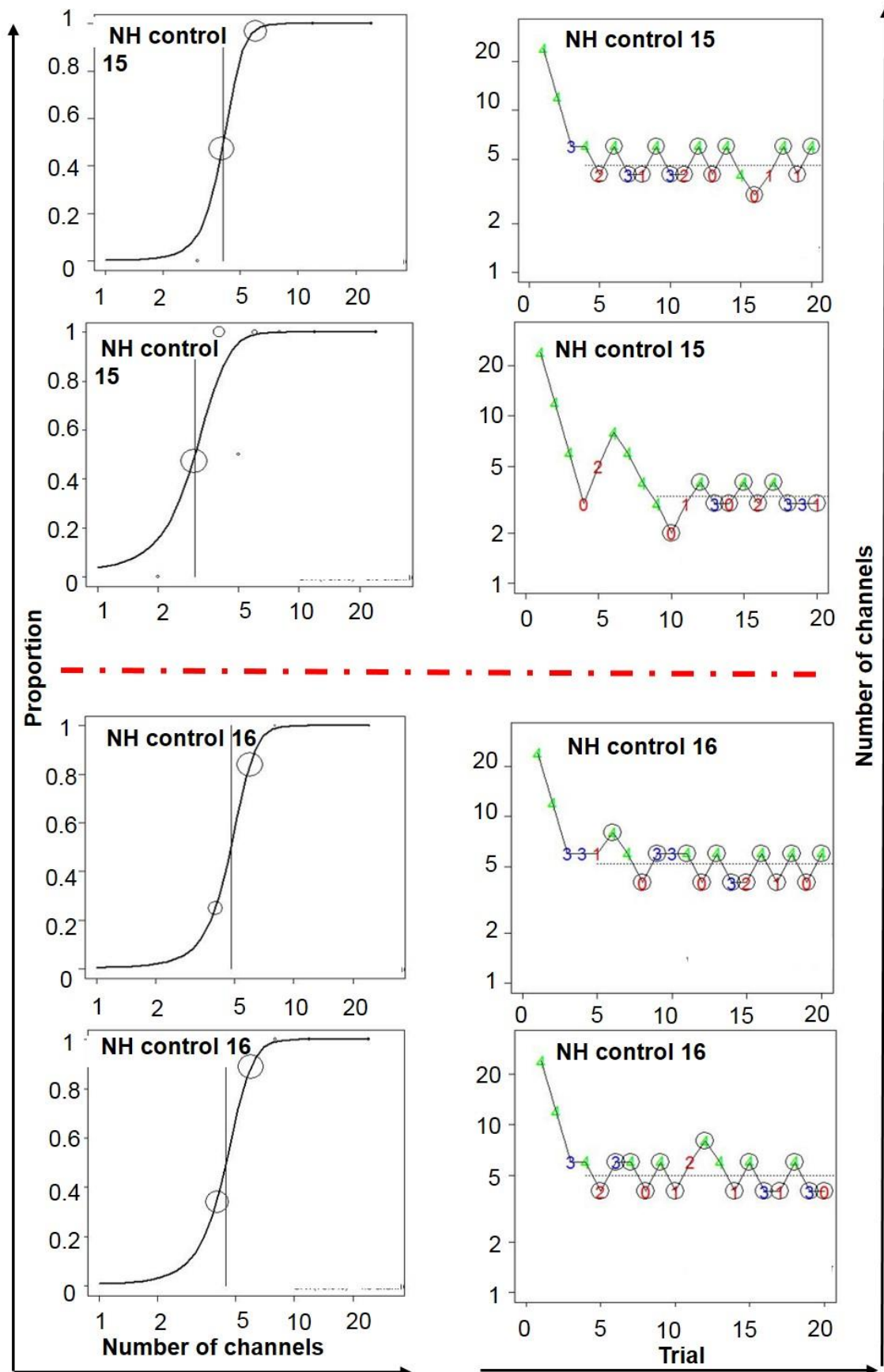
Appendix 5.20: As for Appendix 5.15, but output this time from NH controls 11 & 12 (Experiment 7).



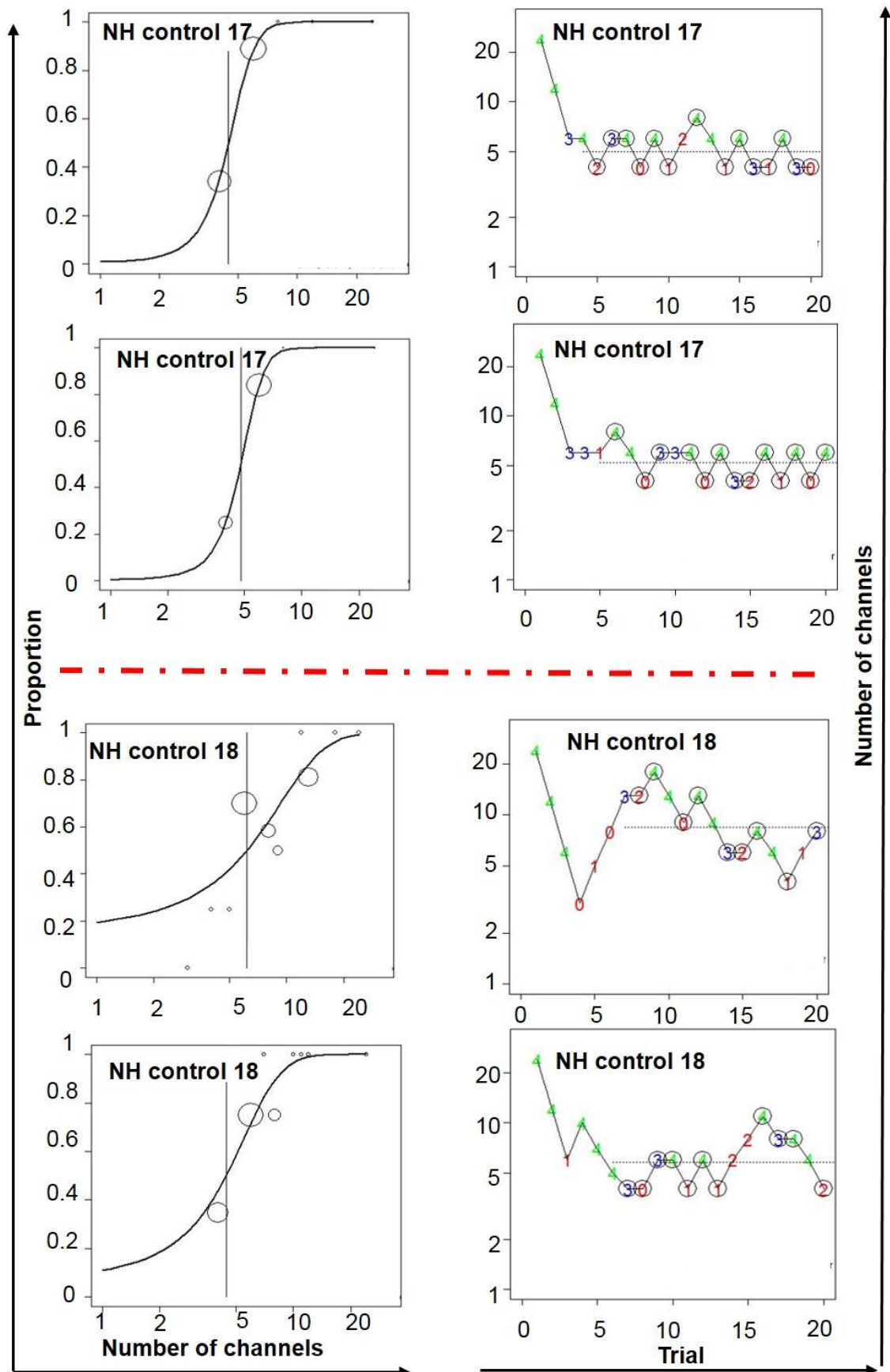
Appendix 5.21: As for Appendix 5.15, but output this time from NH controls 13 & 14 (Experiment 7).



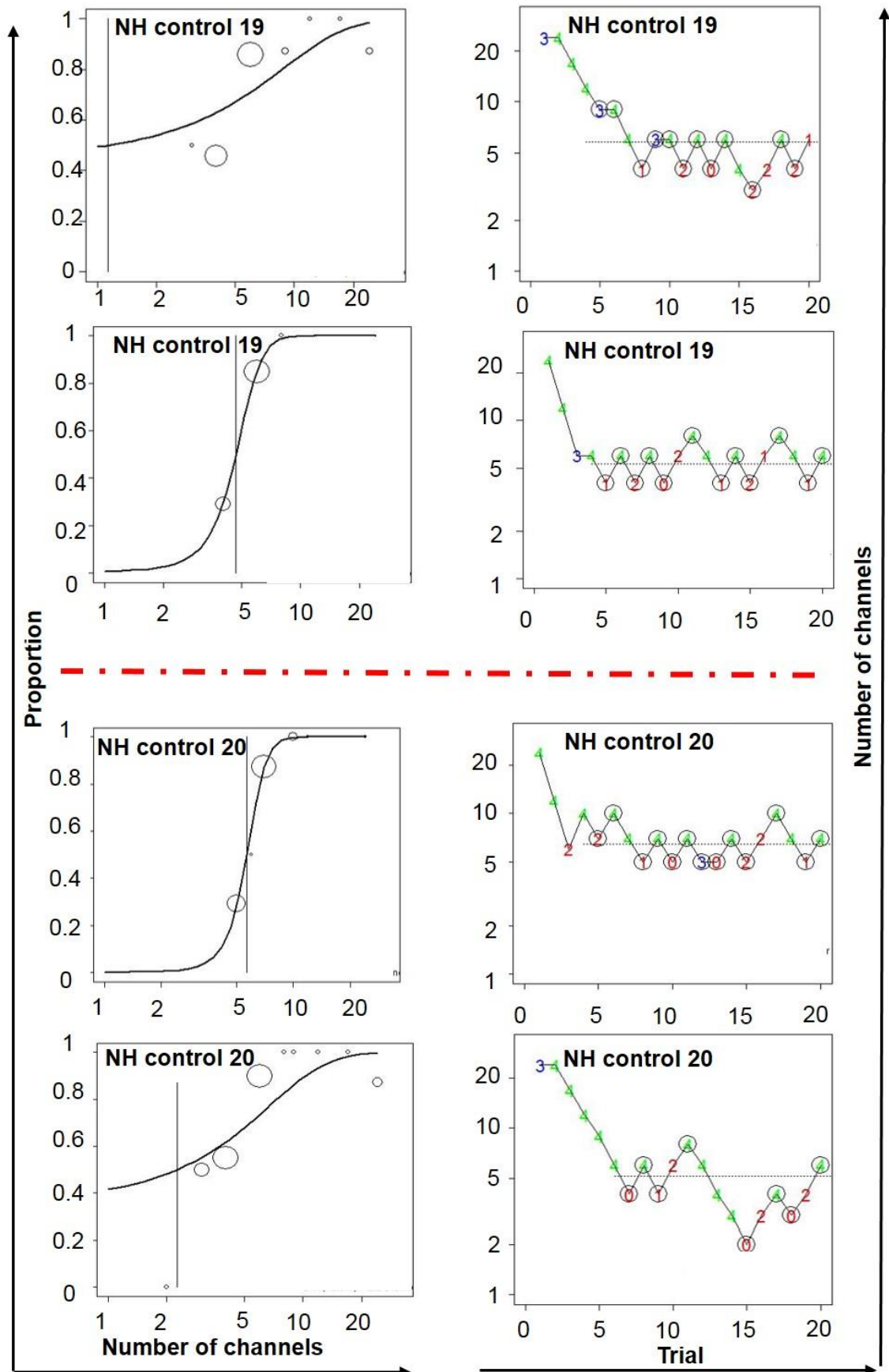
Appendix 5.22: As for Appendix 5.15, but output this time from NH controls 15 & 16 (Experiment 7).



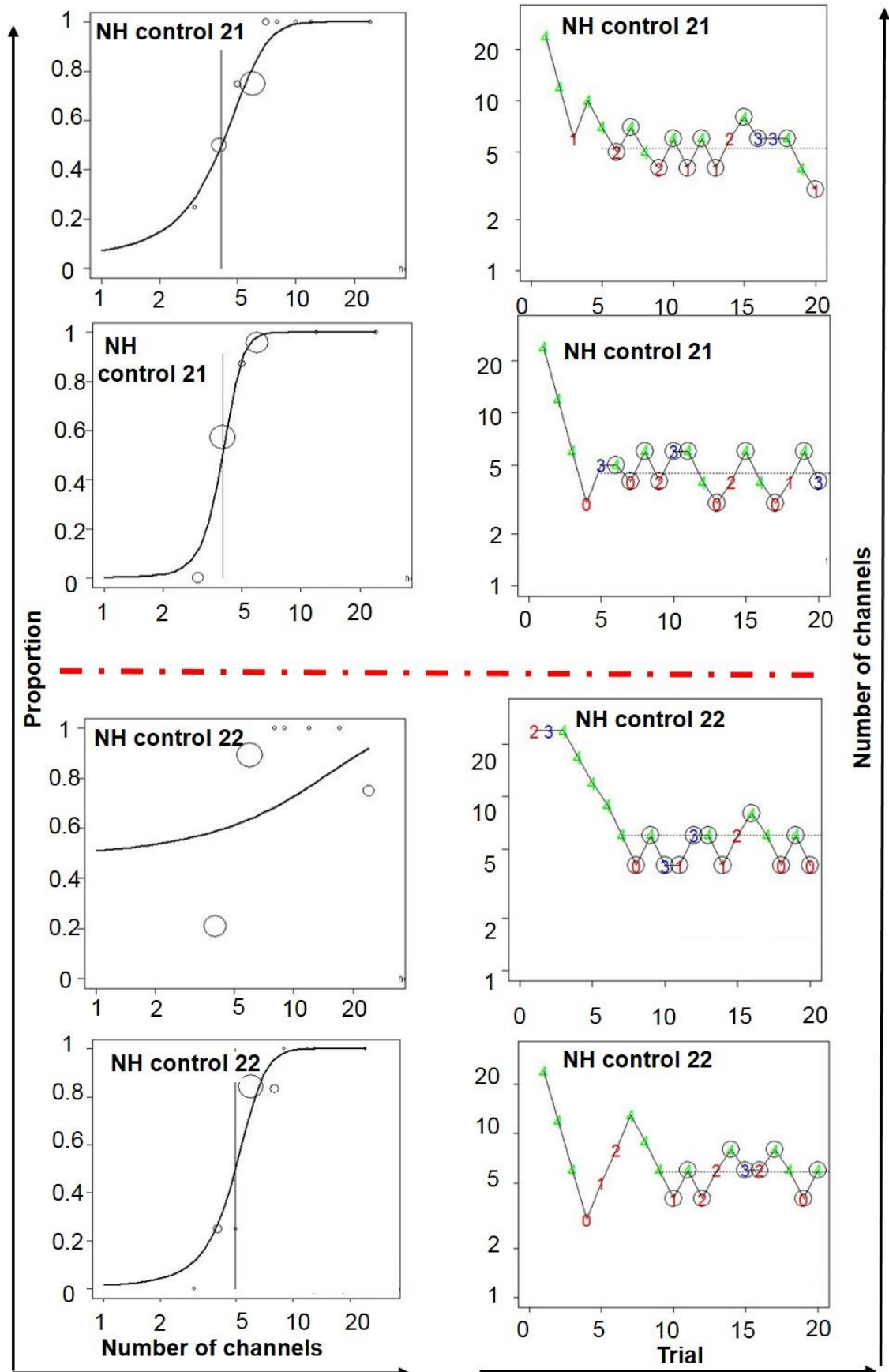
Appendix 5.23: As for Appendix 5.15, but output this time from NH controls 17 & 18 (Experiment 7).



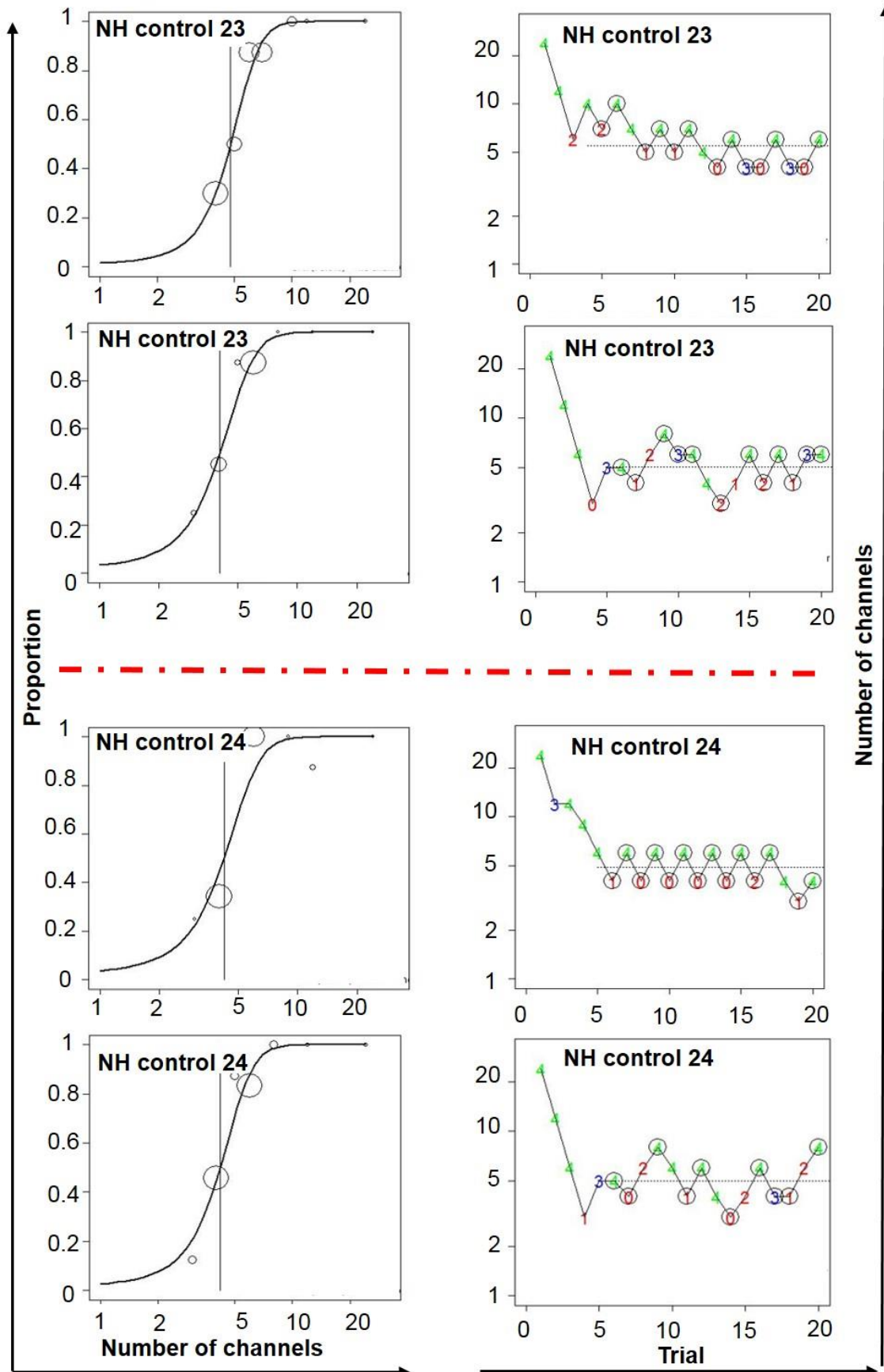
Appendix 5.24: As for Appendix 5.15, but output this time from NH controls 19 & 20 (Experiment 7).



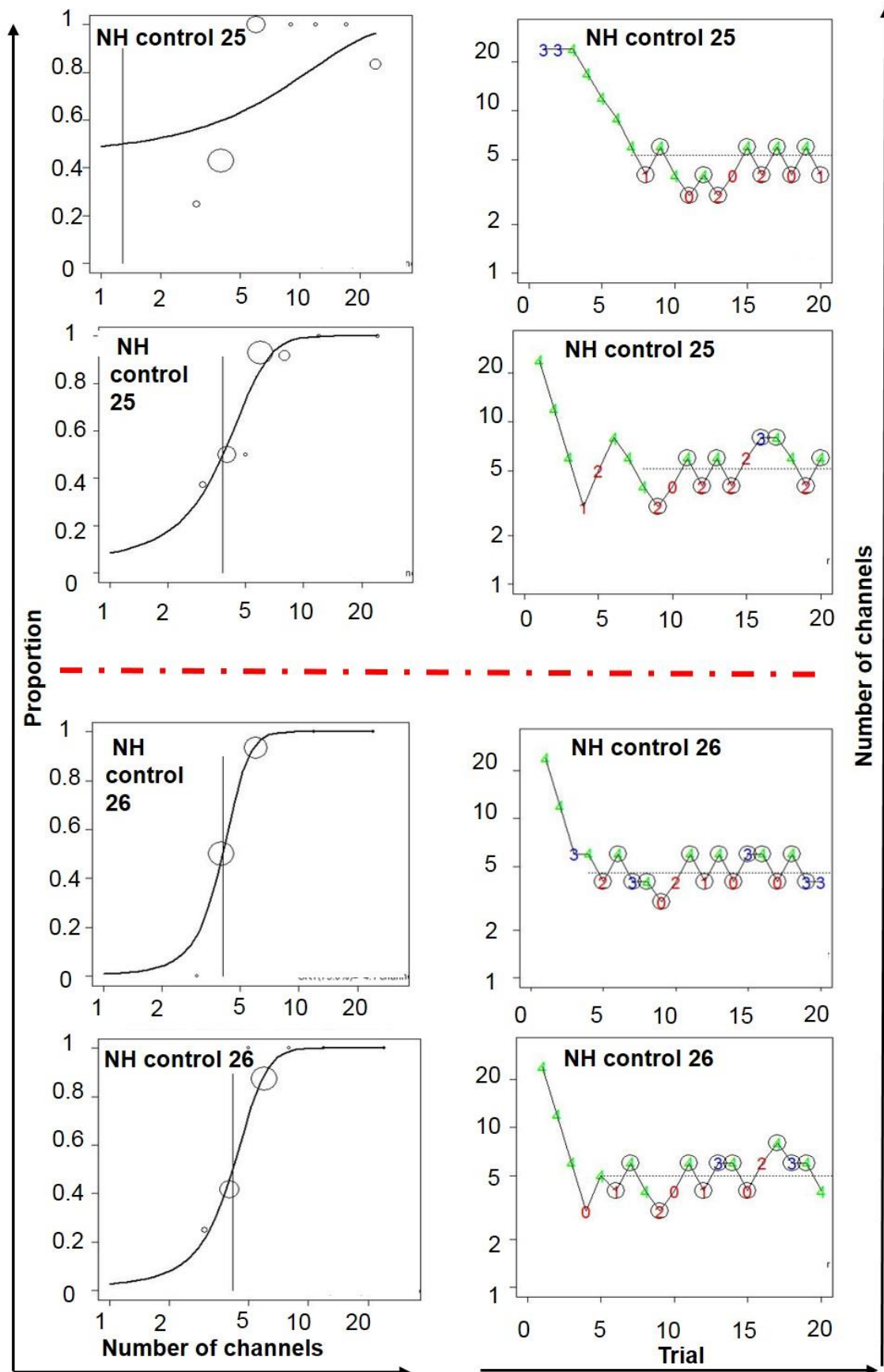
Appendix 5.25: As for Appendix 5.15, but output this time from NH controls 21 & 22 (Experiment 7).



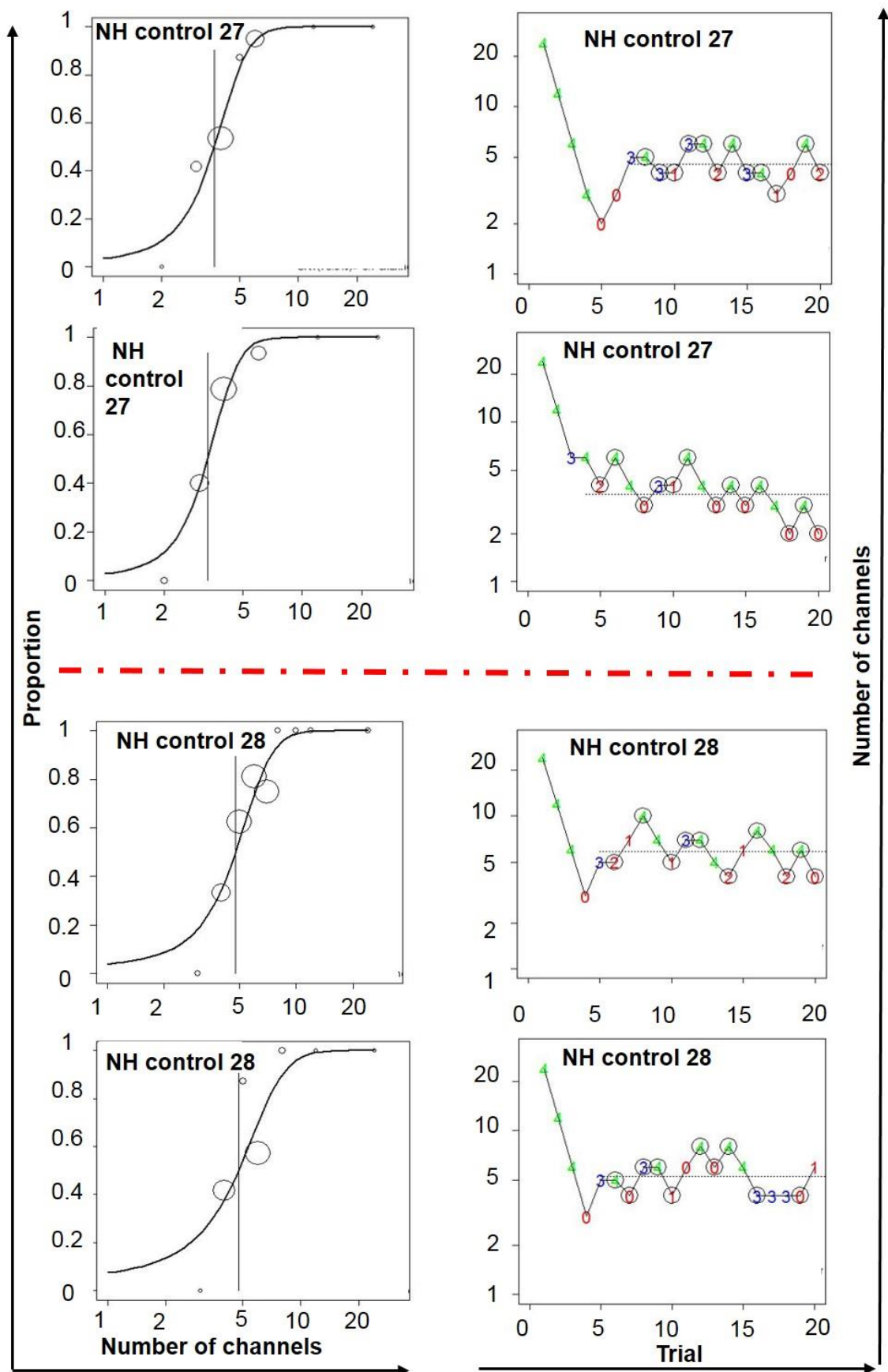
Appendix 5.26: As for Appendix 5.15, but output this time from NH controls 23 & 24 (Experiment 7).



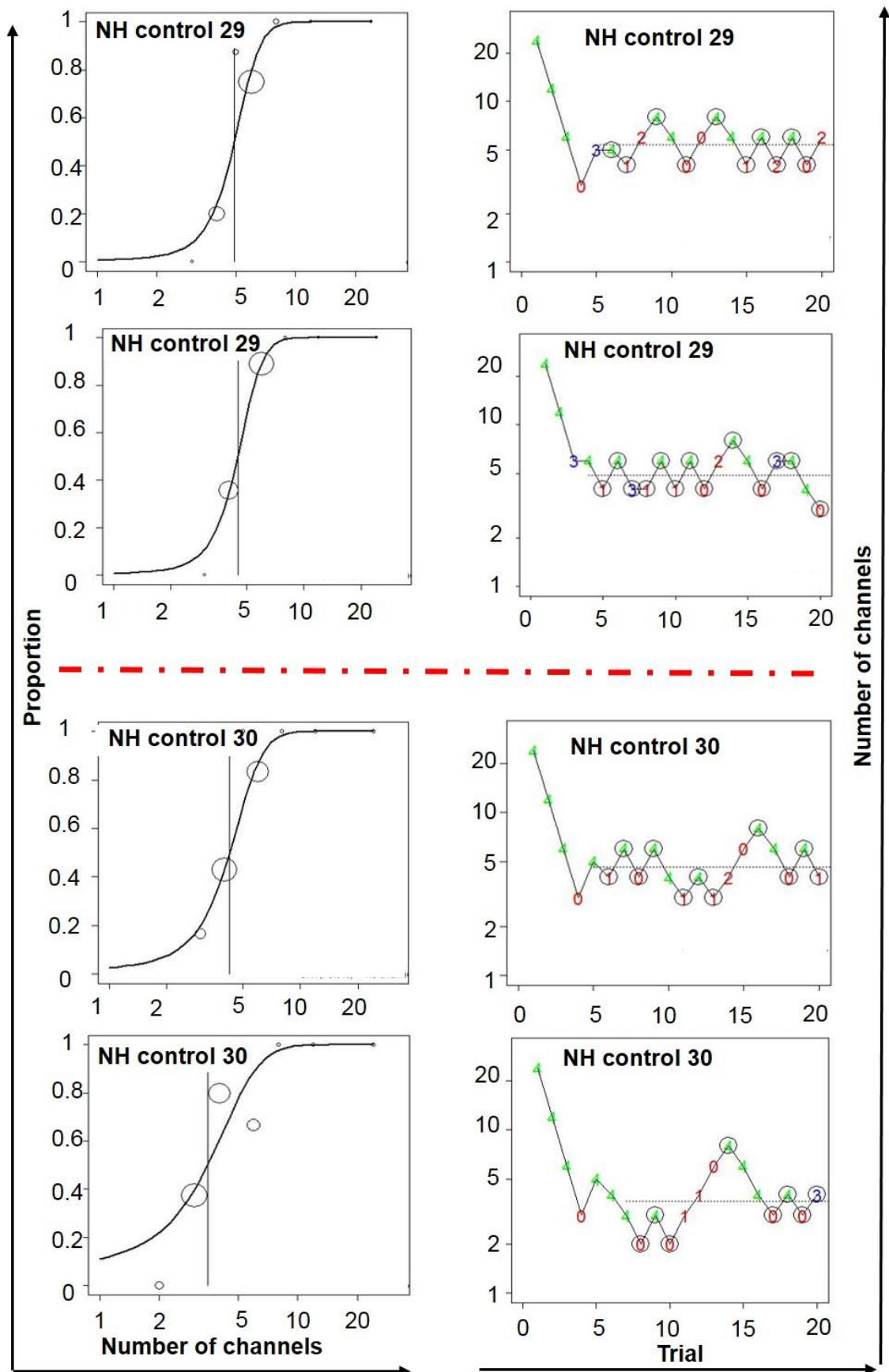
Appendix 5.27: As for Appendix 5.15, but output this time from NH controls 25 & 26 (Experiment 7).



Appendix 5.28: As for Appendix 5.15, but output this time from NH controls 27 & 28 (Experiment 7).

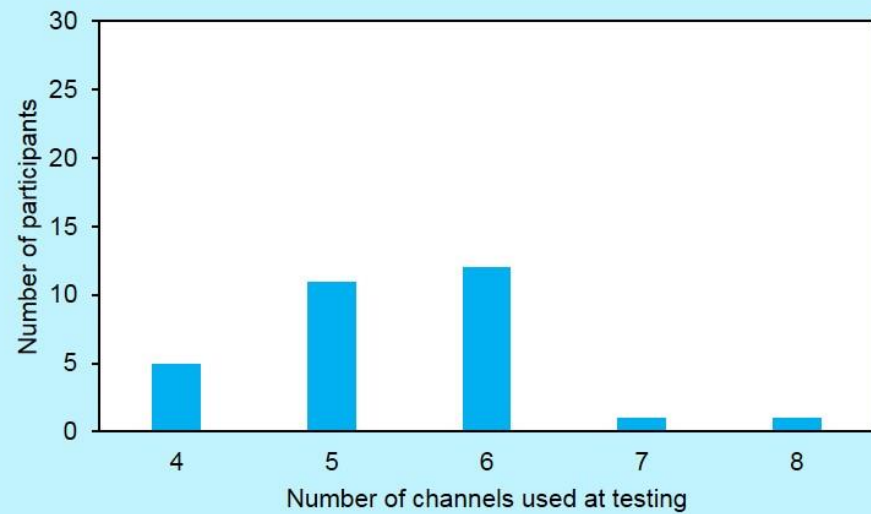


Appendix 5.29: As for Appendix 5.15, but output this time from NH controls 29 & 30 (Experiment 7).

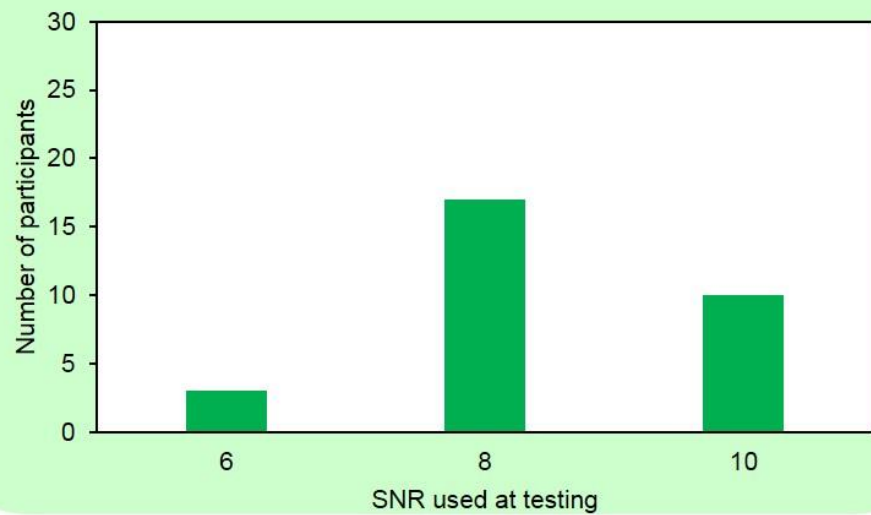


Appendix 5.30: Breakdown of how many participants required each number of channels (panel **a**) and noise level (panel **b**) for the vocoding implemented in Experiment 7. Unit of SNR = dB.

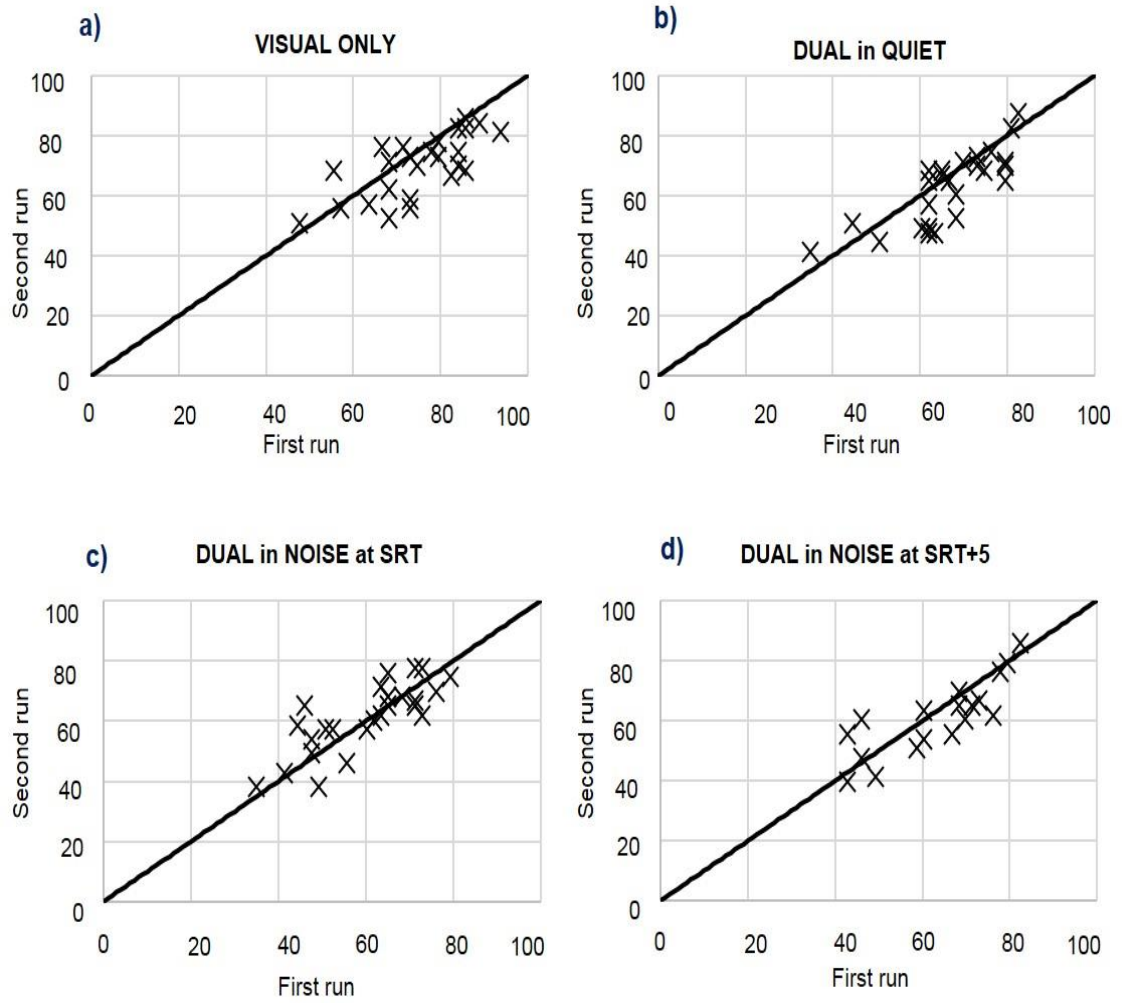
a)



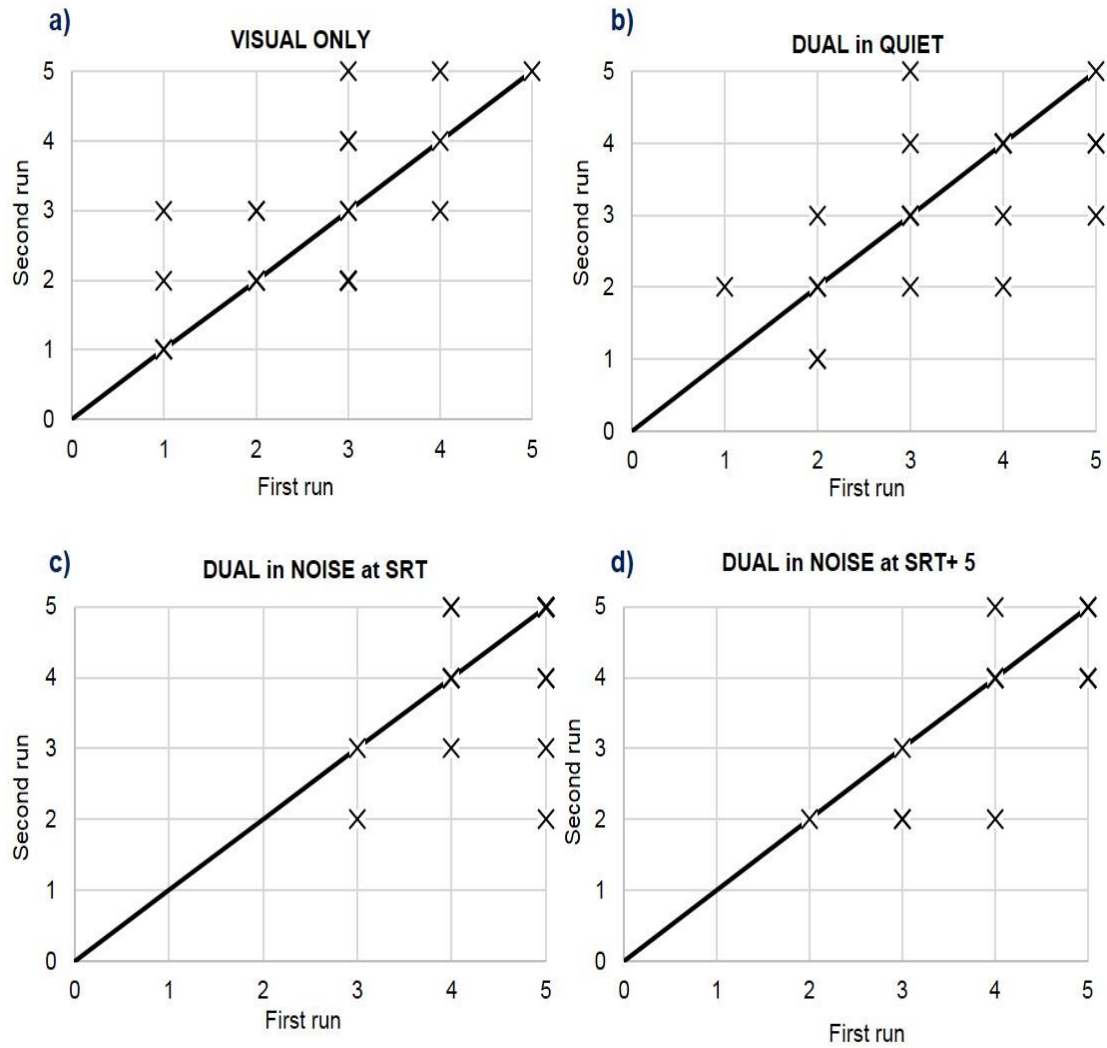
b)



Appendix 6.1: The visual accuracy scores (%) obtained in the two test runs across test conditions of Experiment 8.

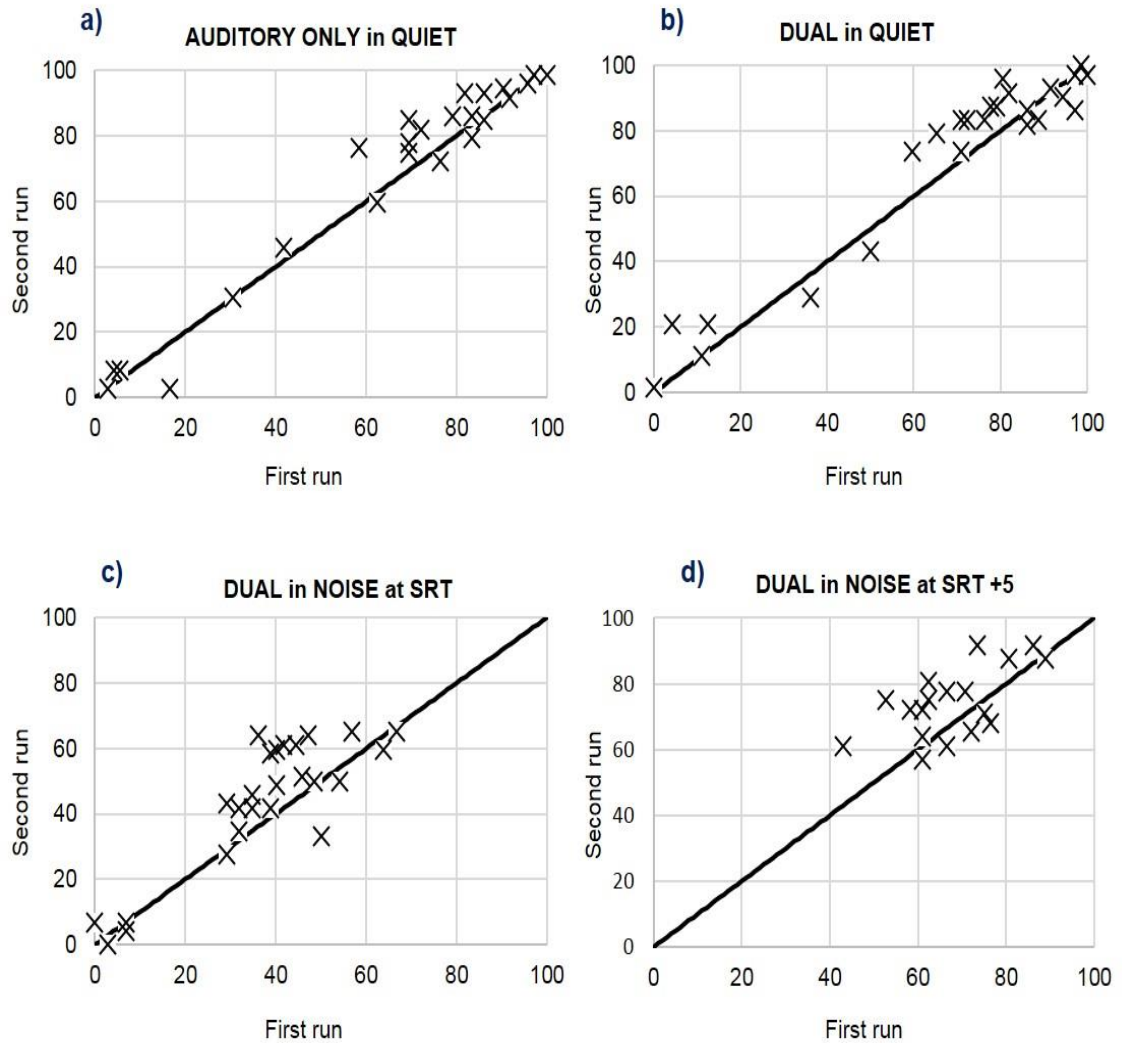


Appendix 6.2: The subjective ratings obtained in the two test runs across test conditions of Experiment 8.

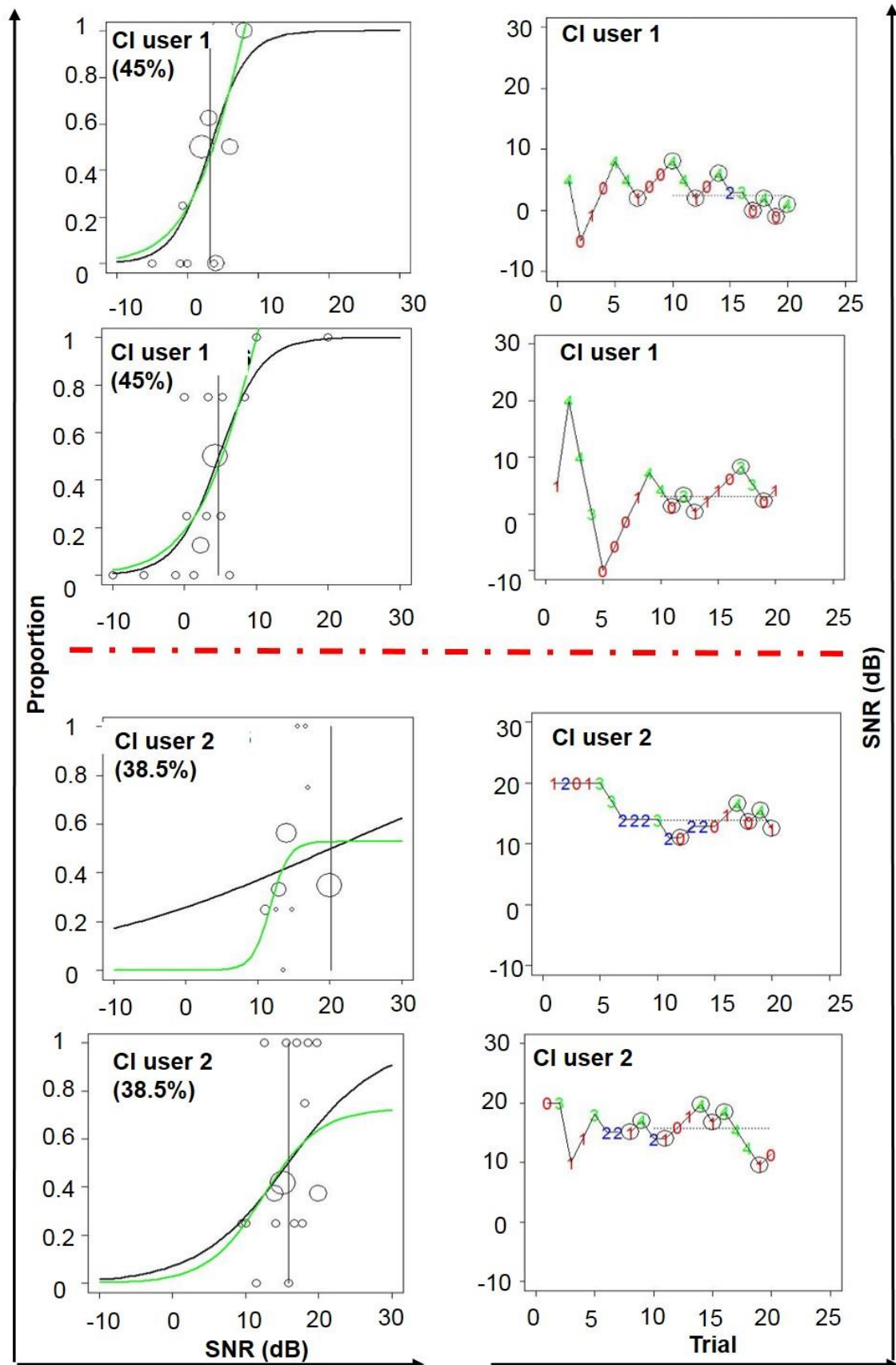


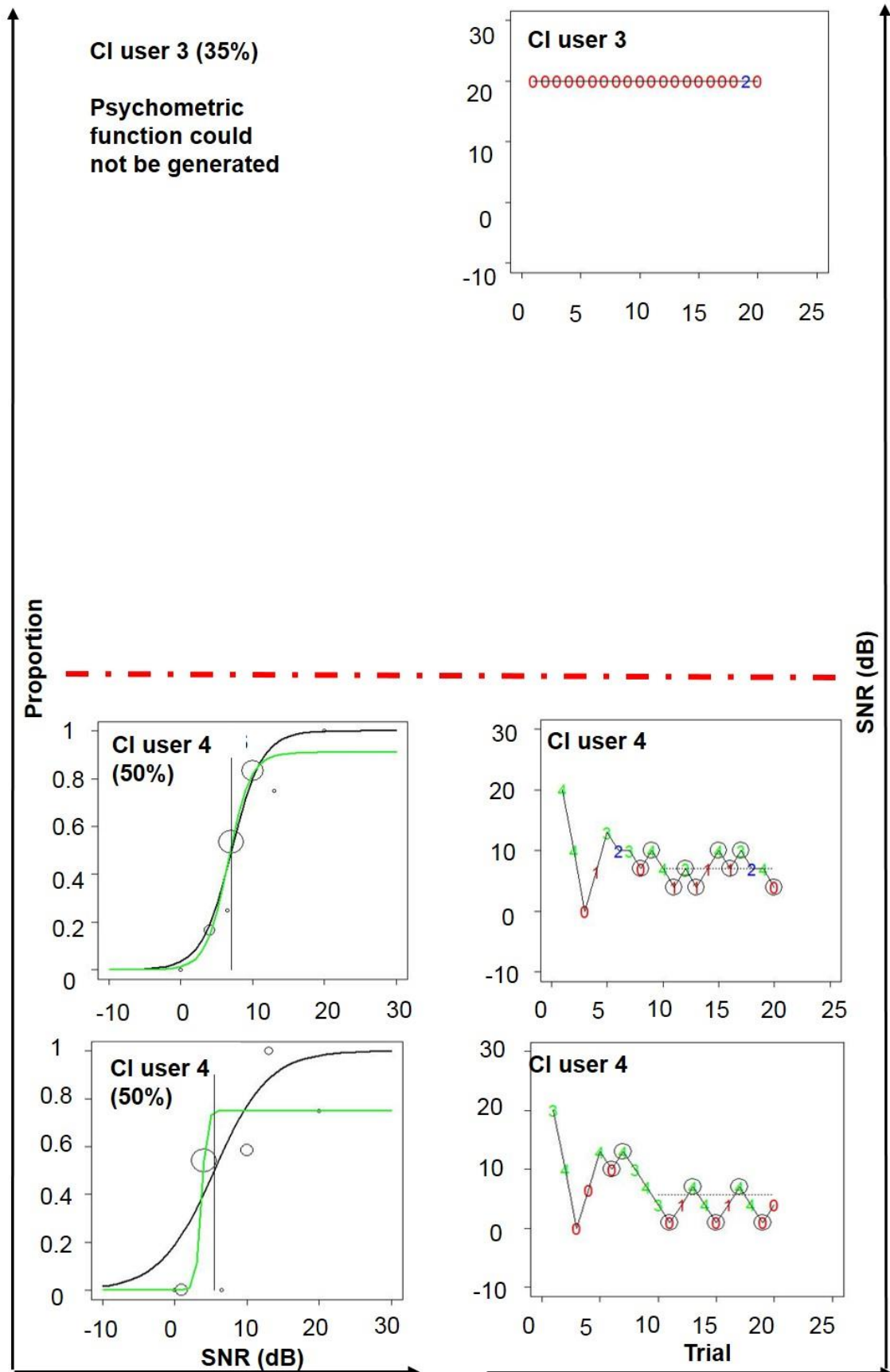
1 = Not at all hard work
 2 = Quite hard work
 3 = Medium hard work
 4 = Very hard work
 5 = Extremely hard work

Appendix 6.3: The auditory accuracy scores (%) obtained in the two test runs across test conditions of Experiment 8.

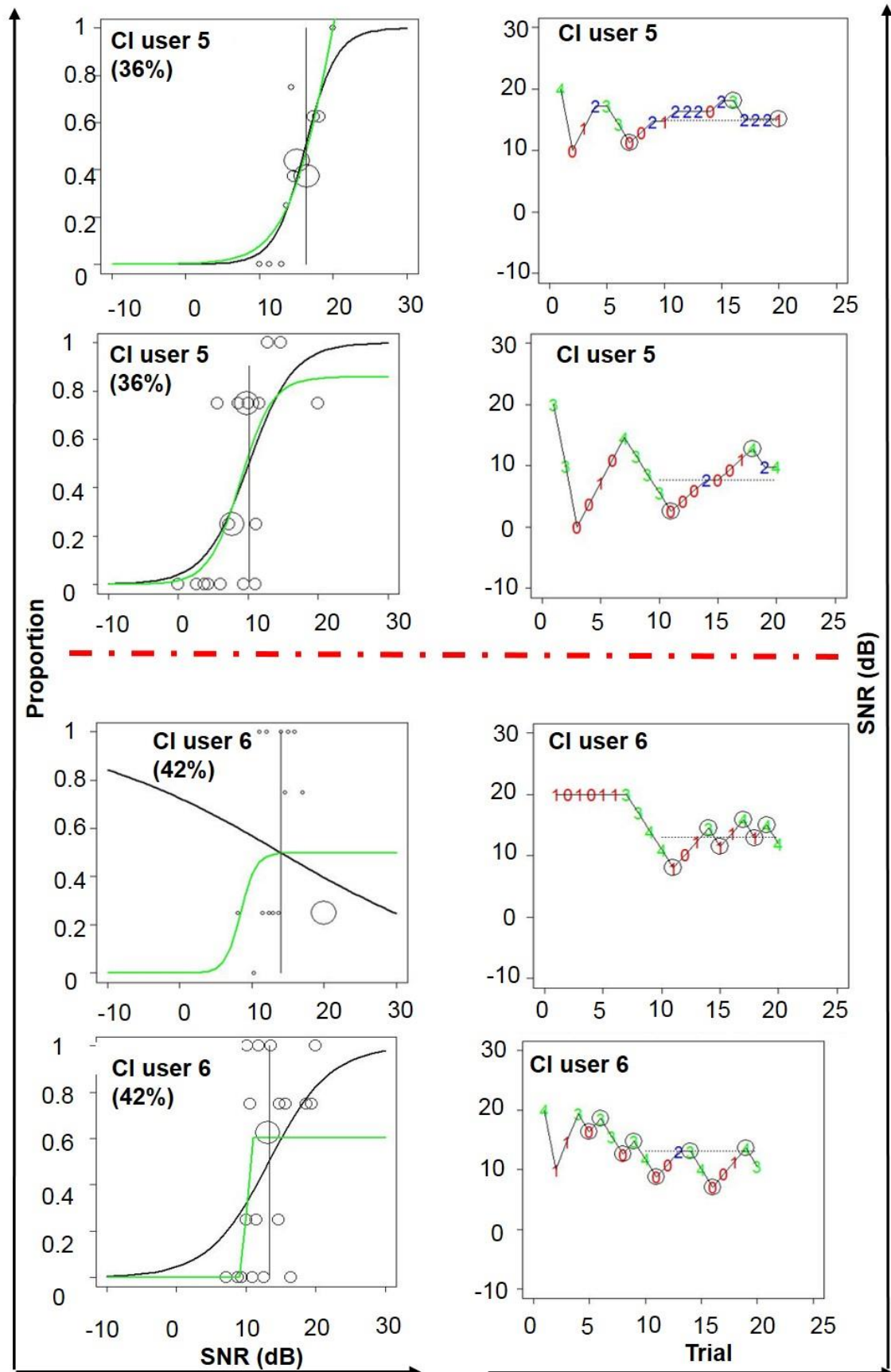


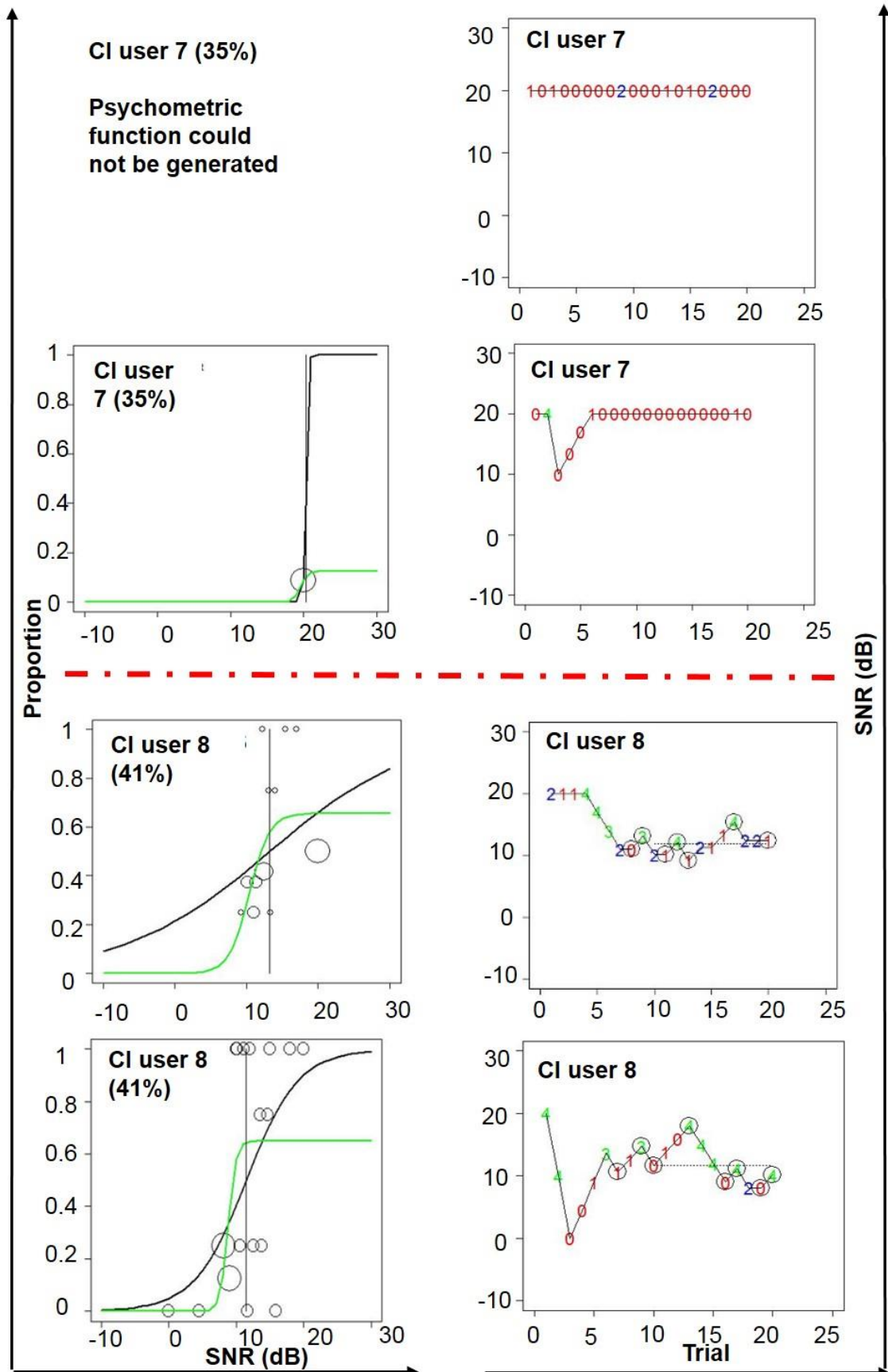
Appendix 6.4: As for Appendix 5.1, but output this time from the adaptive noise tracking of CI users 1 & 2 in both attempts (Experiment 8).

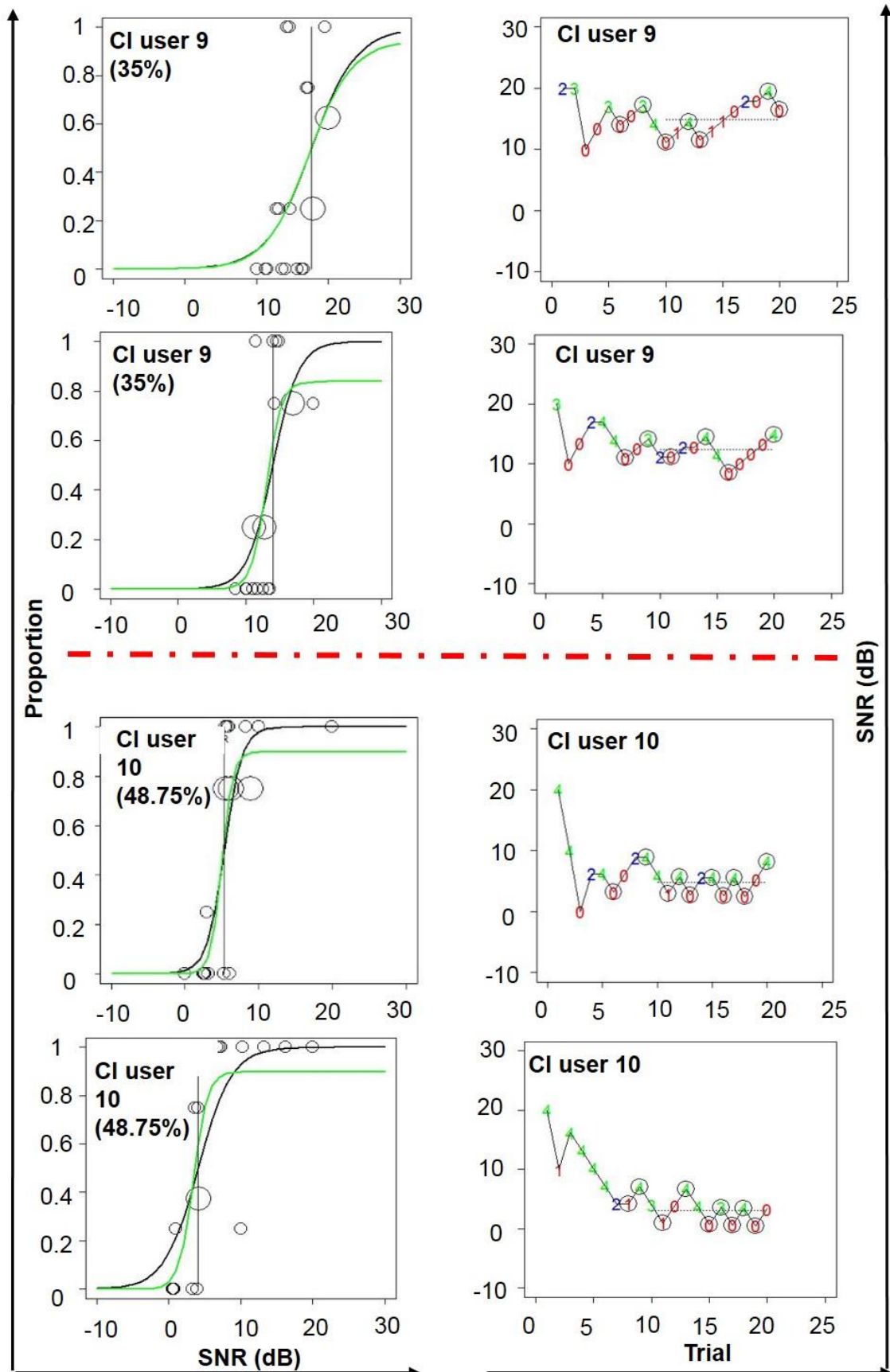


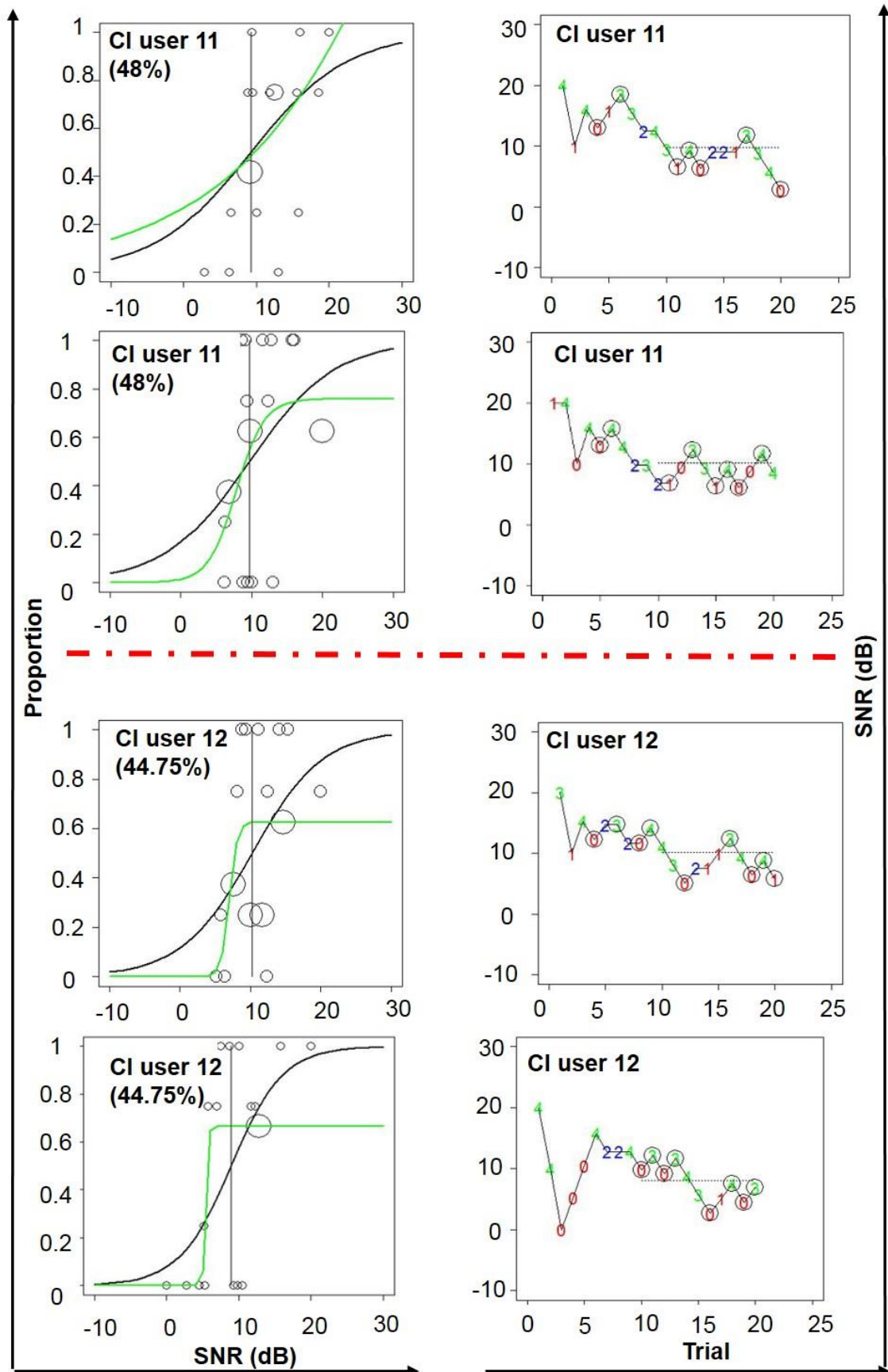
Appendix 6.5: As for Appendix 6.4, but output this time from CI users 3 & 4 (Experiment 8).

Appendix 6.6: As for Appendix 6.4, but output this time from CI users 5 & 6 (Experiment 8).

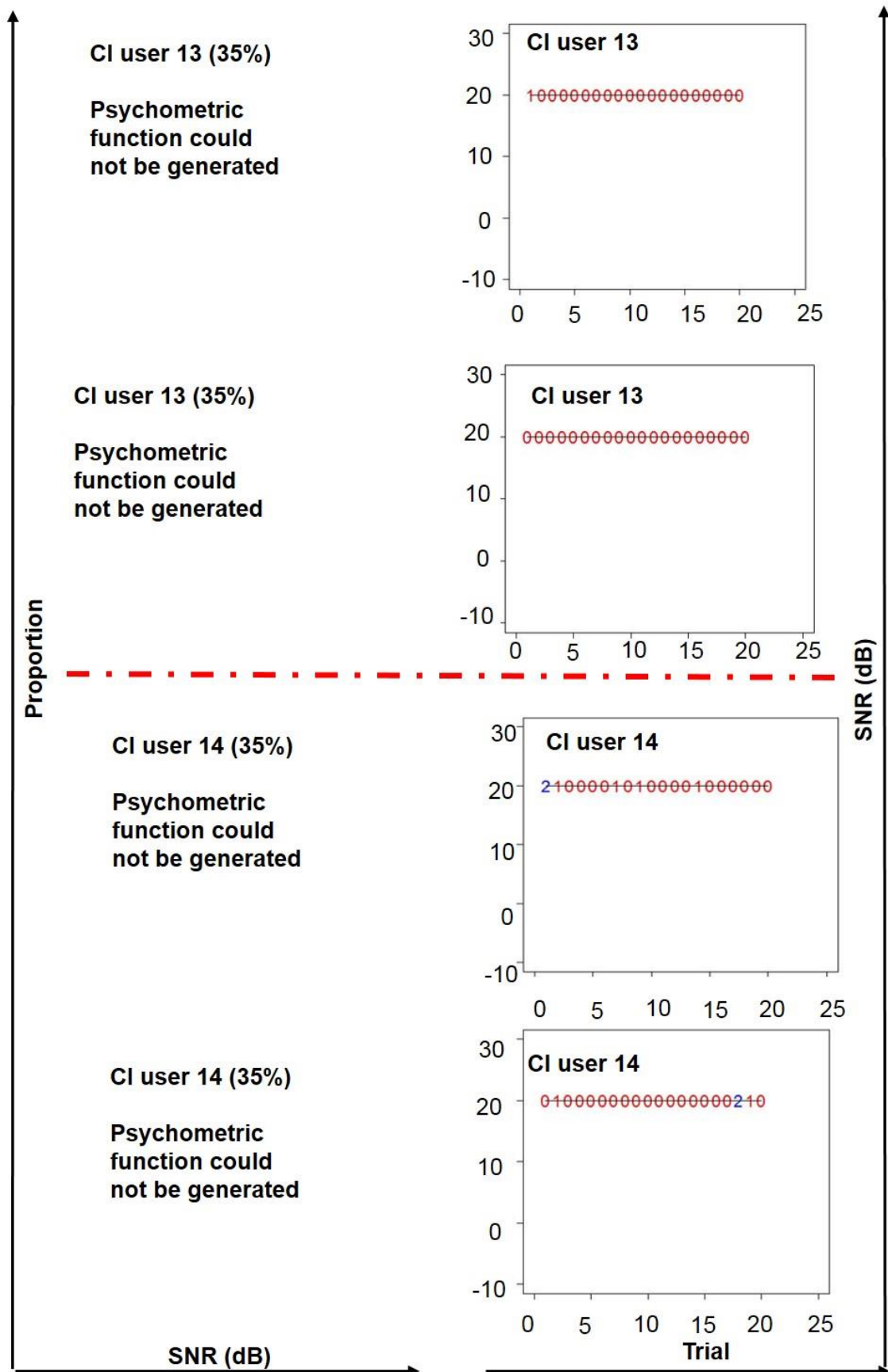


Appendix 6.7: As for Appendix 6.4, but output this time from CI users 7 & 8 (Experiment 8).

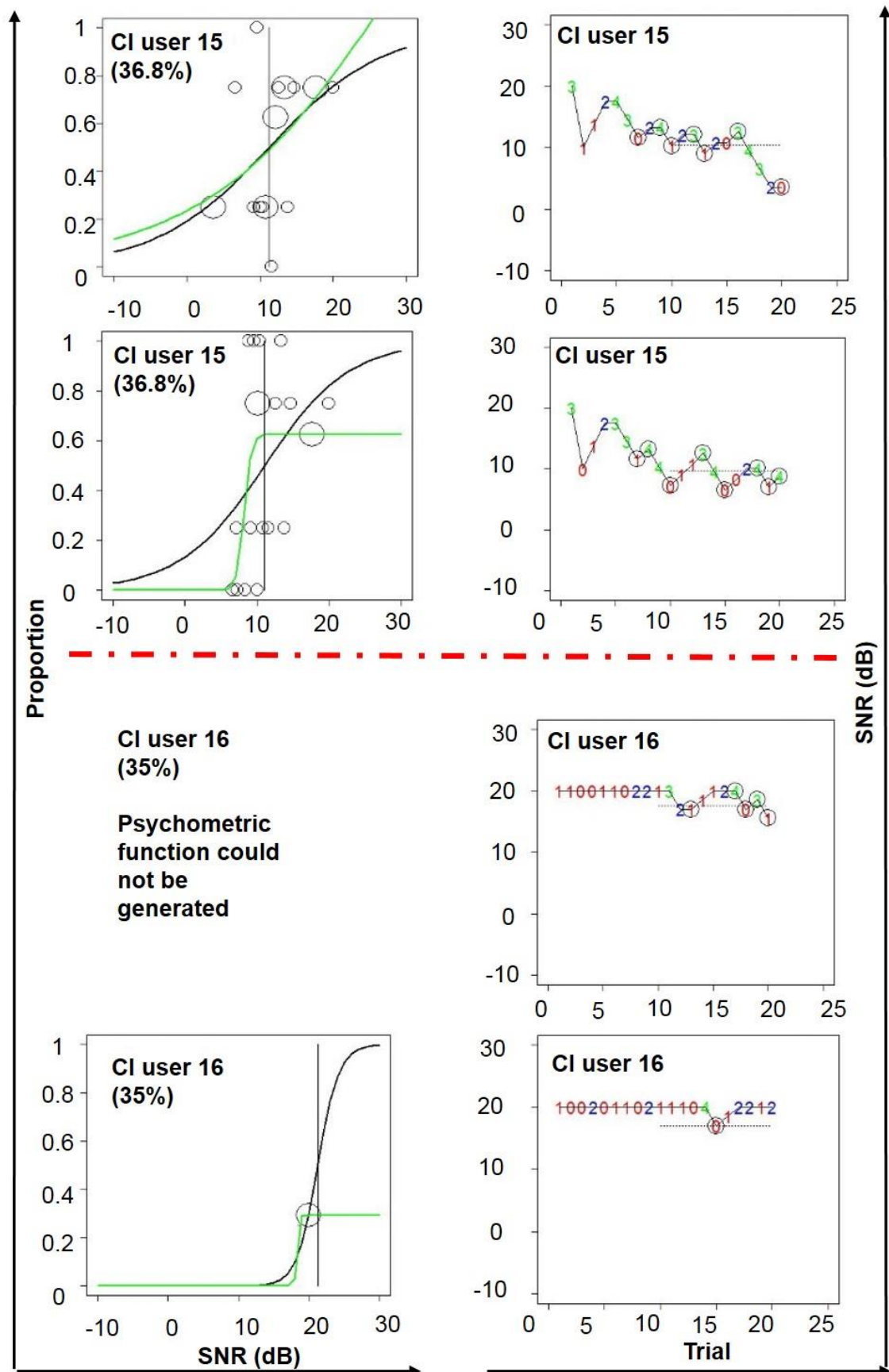
Appendix 6.8: As for Appendix 6.4, but output this time from CI users 9 & 10 (Experiment 8).

Appendix 6.9: As for Appendix 6.4, but output this time from CI users 11 & 12 (Experiment 8).

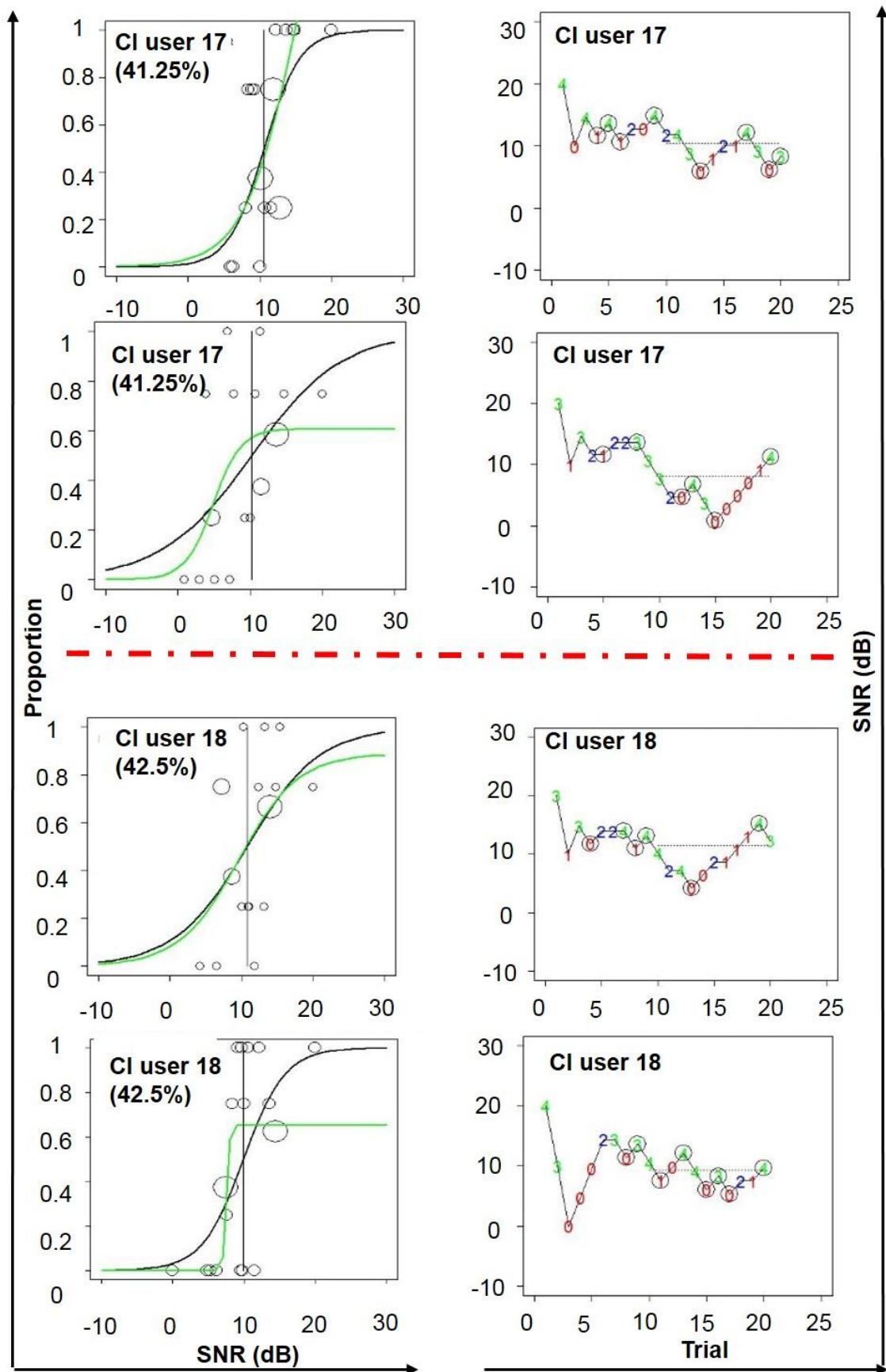
Appendix 6.10: As for Appendix 6.4, but output this time from CI users 13 & 14 (Experiment 8).



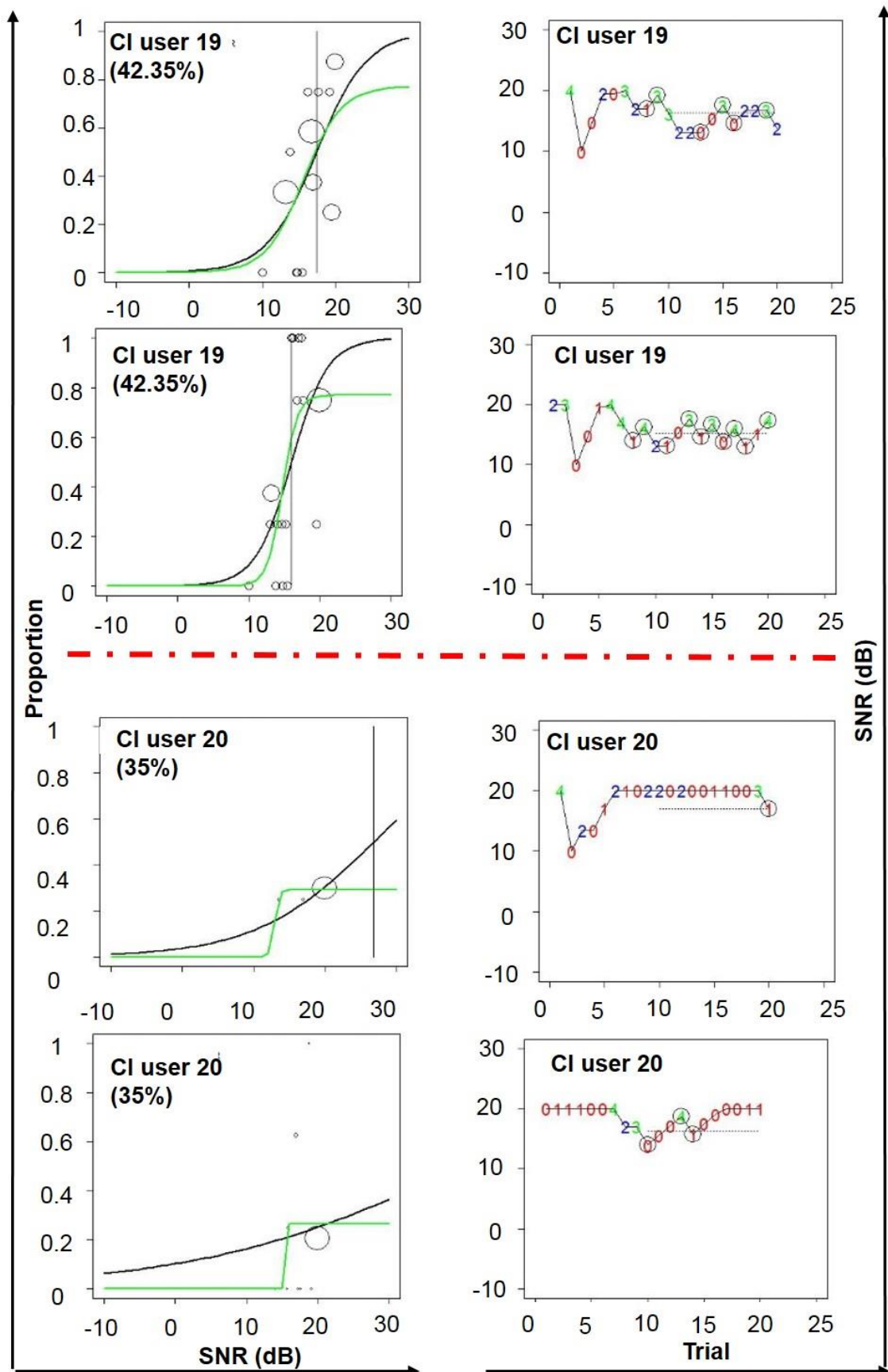
Appendix 6.11: As for Appendix 6.4, but output this time from CI users 15 & 16 (Experiment 8).



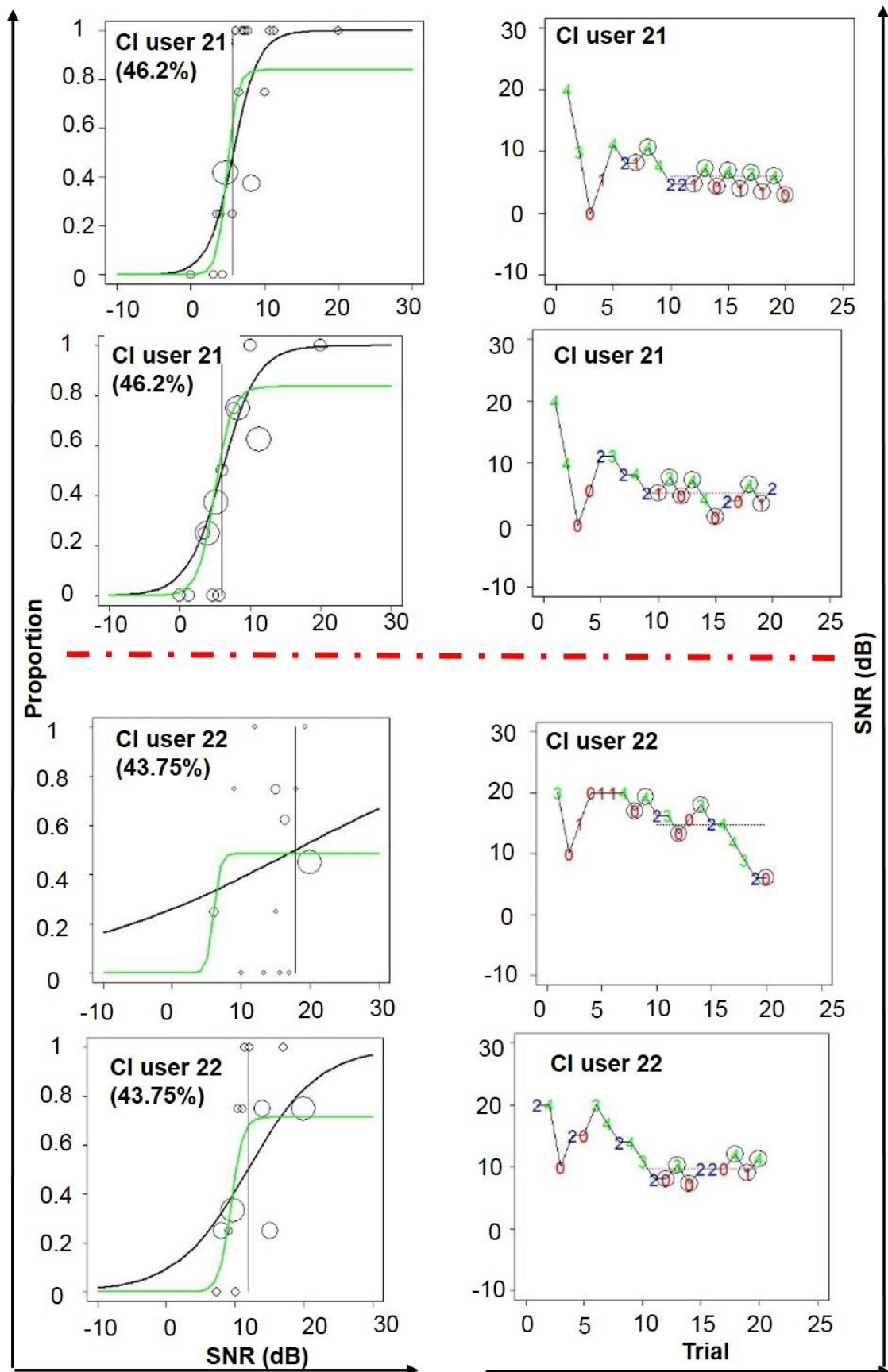
Appendix 6.12: As for Appendix 6.4, but output this time from CI users 17 & 18 (Experiment 8).



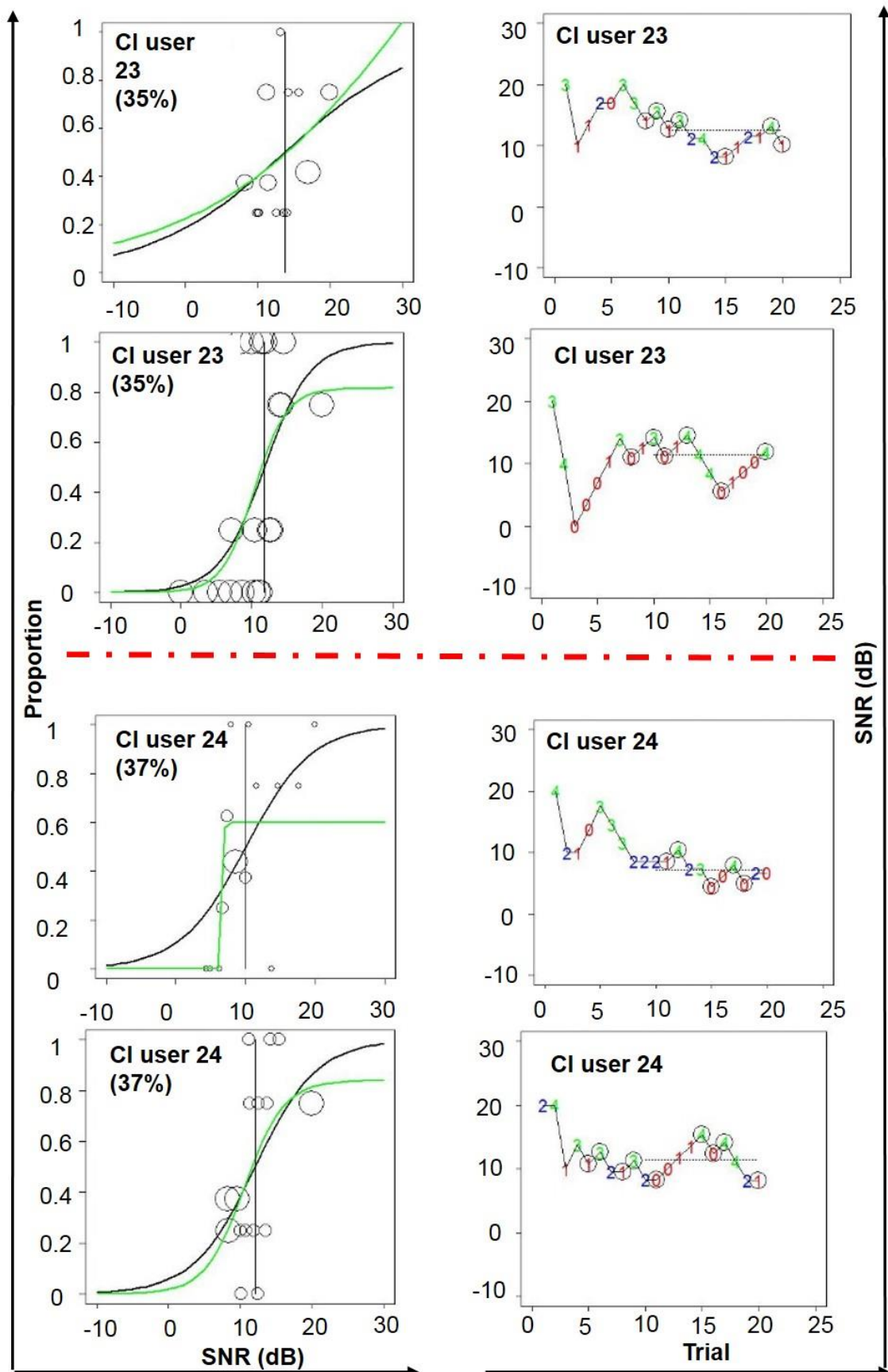
Appendix 6.13: As for Appendix 6.4, but output this time from CI users 19 & 20 (Experiment 8).



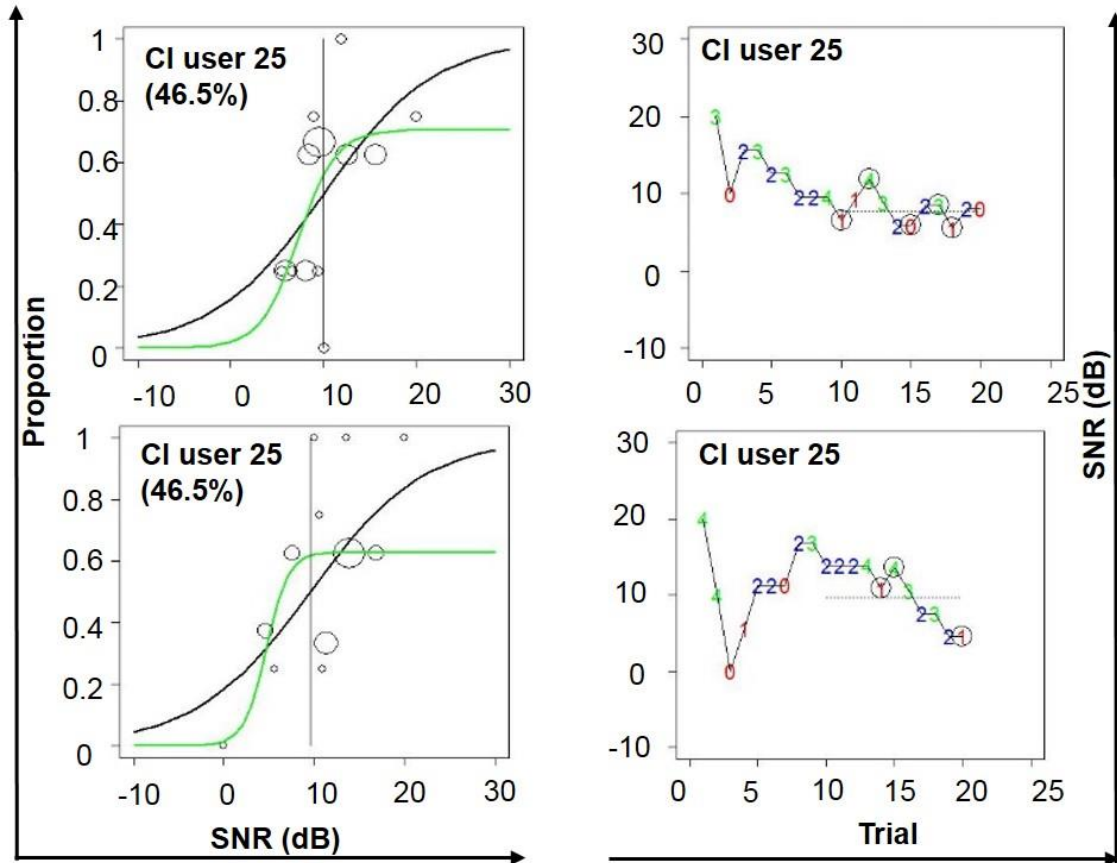
Appendix 6.14: As for Appendix 6.4, but output this time from CI users 21 & 22 (Experiment 8).



Appendix 6.15: As for Appendix 6.4, but output this time from CI users 23 & 24 (Experiment 8).



Appendix 6.16: As for Appendix 6.4, but output this time from CI user 25 (Experiment 8).



REFERENCES

- Alhanbali, S., Dawes, P., Lloyd, S. & Munro, K. J. (2017). Self-reported listening-related effort and fatigue in hearing-impaired adults. *Ear & Hearing*, 38, e39-e48.
- Anderson, C. A., Lazard, D. S. & Hartley, D. E. H. (2017a). Plasticity in bilateral superior temporal cortex: Effect of deafness and cochlear implantation on auditory and visual speech processing. *Hearing Research*, 343, 138-149.
- Anderson, C. A., Wiggins, I. M., Kitterick, P. T. & Hartley, D. E. H. (2017b). Adaptive benefit of cross-modal plasticity following cochlear implantation in deaf adults. *Proceedings of the National Academy of Sciences of the United States of America*, 114, 10256-10261.
- Andreassi, J. L. (2007). *Psychophysiology: Human Behavior and Physiological Response* (5th ed.). Mahway, NJ: Lawrence Erlbaum Associates.
- Arnoldner, C., Riss, D., Brunner, M., Durisin, M., Baumgartner, W. D., Hamzavi, J. S. (2007). Speech and music perception with the new fine structure speech coding strategy: Preliminary results. *Acta Oto-Laryngologica*, 127, 1298-1303.
- Arnsten, A. & Shansky, R. M. (2004). Adolescence: Vulnerable period for stress-induced prefrontal cortical function? Introduction to Part IV. *Annals of the New York Academy of Sciences*, 102, 143-147.
- Aronoff, J. M., Yoon, Y. S., Freed, D. J., Vermiglio, A. J., Pal, I. & Soli, S. D. (2010). The use of interaural time and level difference cues by bilateral cochlear implant users. *Journal of the Acoustical Society of America*, 127, 87-92.
- Aschendorff, A., Kromeier, J., Klenzner, T. & Laszig, R. (2007). Quality control after insertion of the nucleus contour and contour advance electrode in adults. *Ear & Hearing*, 28, 75S-79S.
- ASHA. (2004). Cochlear implants (Technical Report).
- Aston-Jones, G. & Cohen, J. D. (2005). An integrative theory of locus coeruleus-norepinephrine function: Adaptive gain and optimal performance. *Annual Review of Neuroscience*, 28, 403-450.
- Auer, E. T. Jr., Bernstein, L. E., Sungkarat, W. & Singh, M. (2007). Vibrotactile activation of the auditory cortices in deaf versus hearing adults. *Neuroreport*, 18, 645-687.
- AVUK (2016). *What is Auditory Verbal therapy?* Retrieved from <https://www.avuk.org/what-is-auditory-verbal-therapy>.
- Baddeley, A. (1992). Working memory. *Science*, 255, 556-559.
- Baddeley, A. D. & Hitch, G. J. (1974). Working memory. In G. H. Bower (Ed.), *The Psychology of Learning and Motivation* (Vol 8, pp. 47-89). New York, NY: Academic Press.

- Baddeley, A., Logie, R., Nimmo-Smith, I. & Brereton, N. (1985). Components of fluent reading. *Journal of Memory and Language*, 24, 490-502.
- Bandura, A. (1997). *Self Efficacy: The Exercise of Control*. New York, NY: Freeman.
- Barkley, R. A. (2012). *Executive functions: what they are, how they work, and why they evolved*. New York, NY: Guilford Press.
- Başkent, D. (2014). The ear-brain connection in cochlear implant users: learning to listen again. *ENT and Audiology News*, 23, 1-3.
- Başkent, D. & Shannon, R. V. (2005). Interactions between cochlear implant electrode insertion depth and frequency-place mapping. *Journal of the Acoustical Society of America*, 117, 1405-1416.
- Baumann, U. & Nobbe, A. (2004). Pulse rate discrimination with deeply inserted electrode arrays. *Hearing Research*, 196, 49-57.
- Bavelier, D., Dye, M. W. & Hauser, P. C. (2006). Do deaf individuals see better? *Trends in Cognitive Sciences*, 10, 512-518.
- Beatty, J. & Lucero-Wagner, B. (2000). The pupillary system. In J. T. Cacioppo, L. G. Tassinary & G. Berntson (Eds.), *Handbook of Psychophysiology* (pp. 142-162). Cambridge, UK: Cambridge University Press.
- Bernarding, C., Strauss, D. J., Hannemann, R. & Corona-Strauss, F. I. (2012). Quantification of listening effort correlates in the oscillatory eeg activity: a feasibility study. Proceedings from EMBS '12: *IEEE Engineering in Medicine and Biology Society*, pp. 4615-4618.
- Bertoli, S. & Bodmer, D. (2014). Novel sounds as a psychophysiological measure of listening effort in older listeners with and without hearing loss. *Clinical Neurophysiology*, 125, 1030-1041.
- Bertoli, S. & Bodmer, D. (2016). Effects of age and task difficulty on ERP responses to novel sounds presented during a speech-perception-in-noise test. *Clinical Neurophysiology*, 127, 360-368.
- Bess, F. H., Gustafon, S. J., Corbett, B. A., Lambert, W., Camarata, S. M. & Hornsby, B. W. Y. (2016). Salivary cortisol profiles of children with hearing loss. *Ear & Hearing*, 37, 334-344.
- Best, V., Gallun, F. J., Mason, C. R., Kidd Jr, G., Shinn-Cunningham, B. G. (2010). The impact of noise and hearing loss on the processing simultaneous sentences. *Ear & Hearing*, 31, 213-220.
- Bierer, J. A., Bierer, S. M. & Middlebrooks, J. C. (2010). Partial tripolar cochlear implant simulation: spread of excitation and forward masking in the inferior colliculus. *Hearing Research*, 270, 134-142.
- Bizley, J. K. & Cohen, Y. E. (2013). The what, where and how of auditory-object perception. *Nature Reviews Neuroscience*, 14, 693-707.

- Blackwell, H. R. (1953). *Psychophysical thresholds: Experimental studies of methods of measurement* (Rep. No. 36). Ann Arbor: University of Michigan, Bulletin of Engineering Research Institute.
- Blakemore, S. –J. (2008). The social brain in adolescence. *Nature Reviews Neuroscience*, 9, 267-277.
- Blamey, P. J., Pyman, B. C., Clark, G. M., Dowell, R. C., Gordon, M., Brown, A. M. & Hollow, R. D. (1992). Factors predicting postoperative sentence scores in post-linguistically deaf adult cochlear implant users. *Annals of Otology, Rhinology & Laryngology*, 101, 342-348.
- Blamey, P., Artieres, F., Bařkent, D., Bergeron, F., Beynon, A., Burke, E., Dillier, N., Dowell, R., Fraysse, B. & Gallęgo, S. (2013). Factors affecting auditory performance of postlingually deaf adults using cochlear implants: an update with 2251 patients. *Audiology & Neurotology*, 18, 36-47.
- Block, M. L., Zecca, L. & Hong, J. S. (2007). Microglia-mediated neurotoxicity: uncovering the molecular mechanisms. *Nature Reviews Neuroscience*, 8, 57-69.
- Blundon, J.A. & Zakharenko, S. S. (2013). Presynaptic gating of postsynaptic synaptic plasticity: a plasticity filter in the adult auditory cortex. *Neuroscientist*, 19, 465-478.
- Bolia, R., Nelson, W. T., Ericson, M. & Simpson, B. D. (2000). A speech corpus for multitalker communications research, *Journal of the Acoustical Society of America*, 107, 1065-1066.
- Boothroyd, A. & Schauer, A. (2001, October). *Lowest acceptable performance level*. Proceedings from: *Meeting of the International Collegium of Rehabilitative Audiology*. Berkley, CA.
- Boucsein, W. (2012). *Electrodermal Activity* (2nd ed.). New York, NY: Springer.
- Bowers, G., Cullinan, W. E. & Herman, J. P. (1998). Region-specific regulation of glutamimc acid decarboxylase (GAD) mRNA expression in central stress circuits. *Journal of Neuroscience*, 18, 5938-5947.
- Bregman, A. (1990). *Auditory Scene Analysis: The Perceptual Organisation of Sound*. Cambridge, MA: MIT Press.
- Brehm, J. W. & Self, E. A. (1989). The intensity of motivation. *Annual Review of Psychology*, 40, 109-131.
- Breitmeyer, B. G. (2007). Visual masking: past accomplishments, present status, future developments. *Advances in Cognitive Psychology*, 3, 9-20.
- Breitmeyer, B.G., Ro, T., & Öğmen, H. (2004). A comparison of masking by visual and transcranial magnetic stimulation: implications for the study of conscious and unconscious visual processing. *Consciousness & Cognition*, 13, 829–843.

- Breitmeyer, B. G., & Ögmen, H. (2006). *Visual Masking: Time Slices Through Conscious and Unconscious Vision*. Oxford, UK: Oxford University Press.
- Broadbent, D. E. (1971). *Decision and Stress*. London, UK: Academic Press.
- Broadhead, W. E., Kaplan, B. H., James, S. A., Wagner, E. H., Schoenbach, V. J., Grimson, R., Heyden, S., Tibblin, G. & Gehlbach, S. H. (1983). The epidemiologic evidence for a relationship between social support and health. *American Journal of Epidemiology*, 117, 521-537.
- Broadhead, W. E., Gehlbach, S. H., de Gruy, F. V. & Kaplan, B. H. (1988). The Duke-UNC Functional Social Support Questionnaire: Measurement of social support in family medicine patients. *Medical Care*, 26, 706-723.
- Brons, I., Houben, R. & Dreschler, W. A. (2014). Effects of noise reduction on speech intelligibility, perceived listening effort, and personal preference in hearing impaired listeners. *Trends in Hearing*, 1-10.
- Brownley, K. A., Hurwitz, B. E. & Schneiderman, N. (2000). Cardiovascular psychophysiology. In J. T. Cacioppo, L. G. Tassinary & G. G. Berntson (Eds.), *Handbook of Psychophysiology* (2nd ed., pp. 224-264). New York, NY: Cambridge University.
- Brungart, D. S. (2001). Informational and energetic masking effects in the perception of two simultaneous talkers. *Journal of the Acoustical Society of America*, 109, 1101-1109.
- Buckley, D., Codina, C., Bhardwaj, P. & Pascalis, O. (2010). Action video game players and deaf observers have larger Goldmann visual fields. *Vision Research*, 50, 548-556.
- Burmeister, M., McInnis, M. G. & Zöllner, S. (2008). Psychiatric genetics: Progress amid controversy. *Nature Reviews Genetics*, 9, 527-540.
- Byrne, D., Dillon, H., Tran, K., Arlinger, S. & Wibraham, K. (1994). An international comparison of long-term average speech spectra. *Journal of the Acoustical Society of America*, 96, 2108-2120.
- Cabestrero, R., Crespo, A. & Quiros, P. (2009). Pupillary dilation as an index of task demands. *Perceptual and Motor Skills*, 109, 664-678.
- Calandruccio, L. & Smiljanic, R. (2012). New sentence recognition materials developed using a basic non-native English lexicon. *Journal of Speech, Language and Hearing Research*, 55, 1342-1355.
- Caldwell, A. & Nittroer, S. (2013). Speech perception in noise by children with cochlear implants. *Journal of Speech, Language and Hearing Research*, 56, 13-30.
- Cambridge Cognition (2006). *CANTAB Connect Research*. Cambridge, UK: Cambridge Cognition.

- Campbell, F. W. & Robson, J. G. (1968). Application of Fourier analysis to the visibility of gratings. *Journal of Physiology*, 197, 551-566.
- Carvalho Fernando, S., Beblo, T., Schlosser, N., Terfehr, K., Otte, C., Löwe, B. & Wingenfeld, K. (2012). Associations of childhood trauma with hypothalamic-pituitary-adrenal function in borderline personality disorder and major depression. *Psychoneuroendocrinology*, 37, 1659-1668.
- Casey, B. J., Giedd, J. N. & Thomas, K.M. (2000). Structural and functional brain development and its relation to cognitive development. *Biological Psychology*, 54, 241-257.
- Casey, B. J., Thomas, K. M., Davidson, M. C., Kunz, K., Franzen, P. I. (2002). Dissociating striatal and hippocampal function development with a stimulus-response compatibility task. *Journal of Neuroscience*, 22, 8647-8652.
- Casey, B. J., Galvan, A. & Hare, T. A. (2005a). Changes in cerebral functional organization during cognitive development. *Current Opinion in Neurobiology*, 15, 239-244.
- Casey, B. J., Tottenham, N., Liston, C. & Durston, S. (2005b). Imaging the developing brain: what have we learned about cognitive development? *Trends in Cognitive Sciences*, 9, 104-110.
- Casey, B. J., Jones, R. M. & Hare, T. A. (2008). The adolescent brain. *Annals of the New York Academy of Sciences*, 1124, 111-126.
- Champoux, F., Lepore, F., Gagné, J. –P. & Théoret, H. (2009). Visual stimuli can impair auditory processing cochlear implant users. *Neuropsychologia*, 47, 17-22.
- Chan, J. C., Freed, D. J., Vermiglio, A. J. & Soli, S. D. (2008). Evaluation of binaural functions in bilateral cochlear implant users. *International Journal of Audiology*, 47, 296-310.
- Chasteen, A. L., Pichora-Fuller, M. K., Dupuis, K., Smith, S. & Singh, G. (2015). Do negative views of aging influence memory and auditory performance through self-perceived abilities? *Psychology of Aging*, 30, 881-893.
- Chatelain, M. & Gendolla, G. H. E. (2015). Implicit fear and effort-related cardiac response. *Biological Psychology*, 111, 73-82.
- Choi, S., Lotto, A., Lewis, D., Hoover, B. & Stelmachowicz, P. (2008). Attentional modulation of word recognition by children in a dual-task paradigm. *Journal of Speech, Language and Hearing Research*, 51, 1042-1054.
- Chun, S., Bayazitov, I.T., Blundon, J. A. & Zakharenko, S. S. (2013). Thalamocortical long-term potentiation becomes gated after the early critical period in the auditory cortex. *Journal of Neuroscience*, 33, 7345-7357.

- Clark, G. M. (2015). The multi-channel cochlear implant: Multi-disciplinary development of electrical stimulation of the cochlea and the resulting clinical benefit. *Hearing Research*, 322, 4-13.
- Cochlear Americas. (2013). *Capture every moment: Adding science to the art of fitting*. Retrieved from <http://pronews.cochlearamericas.com/capture-every-moment-adding-science-to-the-art-of-fitting/>.
- Cochlear. (2015). *Data logging – an aid to practice in CI programming*. Retrieved from <https://www.cochlear.com/uk/for-professionals/sound-connection/data-logging-an-aid-to-clinical-practice-in-ci-programming-and-counselling>.
- Cochlear. (2017). *Cochlear Implant Comparison*. Retrieved from http://www.cochlear.com/wps/wcm/connect/3b26cda4-7b9a-44c3-8163-66b0fc2a058d/FUN2428+ISS2+JAN17+Product+Comparison+Chart+Final+for+MC+01.03.2017+.pdf?MOD=AJPERES&CONVERT_TO=url&CACHEID=ROOTWORKSPACE-3b26cda4-7b9a-44c3-8163-66b0fc2a058d-IBLMHg.
- Cohen, S. (2004). Social relationships and health. *American Psychologist*, 59, 676-684.
- Cohen, S. & Wills, T. A. (1985). Stress, social support, and the buffering hypothesis. *Psychological Bulletin*, 98, 310-357.
- Cohen, N. L., Waltzmann, S. B., Roland, J. T., Staller, S. J. & Hoffman, R. A. (1999). Early results using the Nucleus CI24M in children. *American Journal of Otology*, 20, 198-204.
- Cohen, L., Manion, L. & Morrison, K. (2003). *Research Methods in Education* (5th ed.). London, UK: Routledge Falmer, Taylor & Francis Group.
- Cohen, S. C., Janicki-Deverts, D. & Miller, G. E. (2007). Psychological stress and disease. *Journal of the American Medical Association*, 298, 1685-1687.
- Cole, M. W., Reynolds, J. R., Power, J. D., Repovs, G., Anticevic, A. & Braver, T. S. (2013). Multi-task connectivity reveals flexible hubs for adaptive task control. *Nature Neuroscience*, 16, 1348-1355.
- Coltheart, M. (1980) The persistence of vision. *Philosophical Transactions of the Royal Society London Biological Sciences*, 8, 57-69.
- Coltheart, V. (1999) (Ed). *Fleeting Memories: Cognition of Brief Visual Stimuli*. Cambridge, MA: MIT Press.
- Combs, L.A. & Polich, J. (2006). P3a from auditory white noise stimuli. *Clinical Neurophysiology*, 117, 1106-1112.
- Conway, C.M., Pisoni, D. B. & Kronenberger, W. G. (2009). The importance of sound for cognitive sequencing abilities: the auditory scaffolding hypothesis. *Current Directions in Psychological Science*, 18, 275-279.

- Conway, C.M., Karpicke, J., Anaya, E.M., Henning, S.C., Kronenberger, W. G. & Pisoni, D. B. (2011a). Nonverbal cognition in deaf children following cochlear implantation: motor sequencing disturbances mediate language delays. *Developmental Neuropsychology*, 36, 237-254.
- Conway, C.M., Pisoni, D. B., Anaya, E. M., Karpicke, J. & Henning, S. C. (2011b). Implicit sequence learning in deaf children with cochlear implants. *Developmental Science*, 14, 69-82.
- Croghan, N. B. H., Duran, S. I. & Smith, Z. M. (2017). Re-examining the relationship between number of cochlear implant channels and maximal speech intelligibility. *The Journal of the Acoustical Society of America*, 142, e537.
- Cullinan, W. E. (2000). GABA(A) receptor subunit expression within hypophysiotropic CRH neurons: a dual hybridization histochemical study. *The Journal of Comparative Neurology*, 419, 344-341.
- Dallos, P. (1992). The active cochlea. *Journal of Neuroscience*, 12, 4575-4585.
- Damen, G. W. J. A., van den Oever-Goltstein, H. L., Langereis, M. C., Chute, P. M. & Myalonus, E. A. M. (2006). Classroom performance of children with cochlear implants in mainstream education. *Annals of Otology, Rhinology & Laryngology*, 115, 542-552.
- Daneman, M. & Carpenter, P.A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behaviour*, 19, 450-466.
- Daneman, M. & Merikle, P. M. (1996). Working memory and language comprehension: A meta-analysis. *Psychonomic Bulletin & Review*, 3, 422-433.
- Danermark, B. & Gellerstedt, L. C. (2004). Psychosocial work environment, hearing impairment and health. *International Journal of Audiology*, 43, 383-389.
- Davis, M. H. & Johnsrude, I. S. (2007). Hearing speech sounds: Top-down influences on the interface between audition and speech perception. *Hearing Research*, 229, 132-147.
- Davis, M. H., Johnsrude, I. S., Hervais-Adelman, A., Taylor, K. & McGettigan, C. (2005). Lexical information drives perceptual learning of distorted speech: Evidence from the comprehension of noise-vocoded sentences. *Journal of Experimental Psychology: General*, 134, 222-241.
- Davidson, L. S., Geers, A. E., Blamey, P. J., Tobey, E. A., Brenner, C. A. (2011). Factors contributing to speech perception scores in long-term paediatric cochlear implant users. *Ear & Hearing*, 32, 19S-26S.
- Dawes, P., Munro, K., Kalluri, S. & Edwards, B. (2014). Acclimatization to hearing aids. *Ear & Hearing*, 35, 203-212.

- Deary, I. J. (1994). Intelligence and auditory discrimination: Separating processing speed and fidelity of stimulus representation. *Intelligence*, 18, 189-213.
- Deary, I. J., Caryl, P. G., Egan, V. & Wright, D. (1989). Visual and auditory inspection time: Their interrelationship and correlations with IQ in high ability subjects. *Personal Individual Differences*, 10, 525-533.
- Defenderfer, J., Kerr-German, A., Hedrick, M. & Buss, A. T. (2017). Investigating the role of temporal lobe activation in speech perception accuracy with normal hearing adults: An event-related fNIRS study. *Neuropsychologia*, 106, 31-41.
- Del Rio-Hortega, P. (1919). El tercer element de los centros nerviosos. I. La microglia en estado normal. II. Intervencion de la microglia en los procesos patologicos. III. Naturaleza probable de la microglia. *Bollettino del la Sociedad Espanola de Biologica*, 9, 69-120.
- Deluca, J. (2005). *Fatigue as a window to the brain*. Cambridge, MA: The MIT Press.
- Demorest, M. E. & Erdman, S. A. (1987). Development of the communication profile for the hearing impaired. *Journal of Speech and Hearing Disorders*, 52, 129-143.
- Denes, P. & Pinson, E. (1963). *The speech chain*. New York, NY: Bell Telephone Laboratories, Inc (Williams & Wilkins Co, Baltimore).
- Desjardins, J. L. & Doherty, K. A. (2013). Age-related changes in listening effort for various types of masker noises. *Ear & Hearing*, 34, 261-272.
- De Vente, W., Olf, M., Van Amsterdam, J. G. C., Kamphuis, J. H. & Emmelkamp, P. M. G. (2003). Physiological differences between burnout patients and healthy controls: Blood pressure, heart rate, and cortisol response. *Occupational and Environmental Medicine*, 60, i54-i61.
- Dhabar, F. S. & McEwen, B. S. (1997). Acute stress enhances while chronic stress suppresses cell-mediated immunity in vivo: a potential role for leukocyte trafficking. *Brain, Behaviour and Immunity*, 11, 286-306.
- Diamond, A. (2013). Executive functions. *Annual Review of Psychology*, 64, 135-168.
- Diamond, D. M., Campbell, A. M., Park, C. R., Halonen, J. & Zoladz, P. R. (2007). The temporal dynamics model of emotional memory processing: A synthesis on the neurobiological basis of stress-induced amnesia, flashbulb and traumatic memories, and the Yerkes-Dodson Law. *Neural Plasticity*, 2007, 1-34.
- Doorduyn, J., De Vries, E. F., Willemsen, A. T., De Groot, J. C., Dierckx, R. A. & Klein, H. C. (2009). Neuroinflammation in schizophrenia-related psychosis: a PET study. *Journal of Nuclear Medicine*, 50, 1801-1807.
- Dorman, M. F., Loizou, P. C. & Rainey, D. (1997). Simulating the effect of cochlear-implant electrode insertion depth on speech understanding. *Journal of the Acoustical Society of America*, 102, 2993-2996.

- Dorman, M. F., Gifford, R. H., Spar, A. J., McKarns, S. A. (2008). The benefits of combining acoustic and electric stimulation for the recognition of speech, voice and melodies. *Audiology & Neurotology*, 13, 105-112.
- Dosenbach, N. U., Fair, D. A., Cohen, A. L., Schlaggar, B. L. & Peterson, S. E. (2008). A dual-networks architecture of top-down control. *Trends in Cognitive Sciences*, 12, 99-105.
- Downs, D. W. (1982). Effects of hearing aid use on speech discrimination and listening effort. *Journal of Speech and Hearing Disorders*, 47, 189-193.
- Drullman, R., Festen, J. M., & Plomp, R. (1994). Effect of reducing slow temporal modulations on speech reception. *The Journal of the Acoustical Society of America*, 95, 2670–2680.
- Dubno, J. R., Dirks, D. D. & Langhofer, L. R. (1982). Evaluation of hearing-impaired listeners using a Nonsense-syllable Test. II. Syllable recognition and consonant confusion patterns. *Journal of Speech and Hearing Research*, 25, 141-148.
- Du Feu, M. & Fergusson, K. (2003). Sensory impairment and mental health. *Advances in Psychiatric Treatment*, 9, 9-103.
- Dye, M. W. (2014). Temporal entrainment of visual attention in children: effects of age and deafness. *Vision Research*, 105, 29-36.
- Dye, M. W. & Bavelier, D. (2010). Attentional enhancements and deficits in deaf populations: an integrative review. *Restorative Neurology and Neuroscience*, 28, 181-192.
- Eastwood, M. R., Corbin, S. L., Reed, M., Nobbs, H. & Kedward, H. B. (1985). Acquired hearing loss and psychiatric illness: an estimate of the prevalence and co-morbidity in a geriatric setting. *British Journal of Psychiatry*, 147, 552-556.
- Eckert, M. A., Teubner-Rhodes, S. & Vaden, K. I. (2016). Is listening in noise worth it? The neurobiology of speech recognition in challenging listening conditions. *Ear & Hearing*, 37, 101S-110S.
- Edeline, J. M. (2012). Beyond traditional approaches to understanding the functional role of neuromodulators in sensory cortices. *Frontiers in Behavioural Neuroscience*, 6, 1-15.
- Edwards, B. (2007). The future of hearing aid technology. *Trends in Amplification*, 11, 31-46.
- Edwards, B. (2016). A model of auditory-cognitive processing and relevance to clinical applicability. *Ear & Hearing*, 37, 85S-91S.
- Engel, G. L. (1977). The need for a new medical model: A challenge for biomedicine. *Science*, 196, 129-136.
- Erber, N. P. (1988). *Communication Therapy for Hearing-Impaired Adults*. Abbotsfield, Australia: Clavis.

- Escudé, B., James, C., Deguine, O., Cochard, N., Eter, E. & Fraysse, B. (2006). The size of the cochlea and predictions of insertion depth angles for cochlear implant electrodes. *Audiology & Neurotology*, 11, 27-33.
- Faravelli, C. & Pallanti, S. (1989). Recent life events and panic disorder. *American Journal of Psychiatry*, 146, 622-626.
- Faulkner, A., Rosen, S. & Smith, C. (2000). Effects of the salience of pitch and periodicity information on the intelligibility of four-channel vocoded speech: Implications for cochlear implants. *Journal of the Acoustical Society of America*, 108, 1877-1887.
- Faulkner, A., Rosen, S. & Green, T. (2012). Comparing live to recorded speech in training the perception of spectrally shifted noise-vocoded speech. *Journal of the Acoustical Society of America*, 132, el336-el342.
- Faulkner, K. F. & Pisoni, D. B. (2013). Some observations about cochlear implants: Challenges and future directions. *Neuroscience Discovery, Herbert Open Access Journals*, 1-10.
- Fayad, J. N. & Linthicum, F. H. Jr. (2006). Multichannel cochlear implants: Relation of histopathology to performance. *Laryngoscope*, 116, 1130-1320.
- Fechner, G. T. (1860). *Elemente der Psychophysik* (2nd ed.) Leipzig, Germany: Breitkopf & Härtel.
- Feuerstein, J. F. (1992). Monaural versus binaural hearing: Ease of listening, word recognition and attentional effort, *Ear & Hearing*, 13, 80-86.
- Figueras, B., Edwards, L. & Langdon, D. (2008). Executive function and language in deaf children. *Journal of Deaf Studies and Deaf Education*, 13, 362-377.
- Fine, I., Finney, E. M., Boynton, G.M. & Dobkins, K. R. (2005). Comparing the effects of auditory deprivation and sign language within the auditory and visual cortex. *Journal of Cognitive Neuroscience*, 17, 1621-1637.
- Finlay-Jones, R. & Brown, G. W. (1981). Types of stressful life events and the onset of anxiety and depressive disorders. *Psychological Medicine*, 11, 803-815.
- Finley, C. C. & Skinner, M. W. (2008). Role of electrode placement as a contributor to variability in cochlear implant outcomes. *Otology & Neurotology*, 29, 920-928.
- Finney, E. M., Fine, I., Dobkins, K. R. (2001). Visual stimuli activate auditory cortex in the deaf. *Nature Neuroscience*, 4, 1171-1173.
- Firszt, J. B., Koch, D. B., Downing, M., Litvak, L. (2007). Current steering creates additional pitch percepts in adult cochlear implant recipients. *Otology & Neurotology*, 28, 629-636.

- Fishman, K. E., Shannon, R. V. & Slattery, W. H. (1997). Speech recognition as a function of the number of electrodes used in the SPEAK cochlear implant processor. *Journal of Speech, Language and Hearing Research*, 40, 1201-1215.
- Flynn, M. C. (2005). Datalogging: A new paradigm in the hearing instrument fitting process. *Hearing Review*, 12, 52-57.
- Foo, C., Rudner, M., Rönnberg, & Lunner, T. (2007). Recognition of speech in noise with new hearing instrument compression release settings requires explicit cognitive storage and processing capacity. *Journal of the American Academy of Audiology*, 18, 618-631.
- Frankish, C. J., Green, L. W., Ratner, P. A., Chomik, T. & Larsen, C. (1996). *Health impact assessment as a tool for the health promotion and public policy: A report submitted to the health promotion development division of Health Canada*. Retrieved from <http://catalogue.iugm.qc.ca/GEIDFile/healthimpact.PDF?Archive=192469191064>
- Fraser, S., Gagné, J. P., Alepins, M. & Dubois, P. (2010). Evaluating the effort expended to understand speech in noise using a dual-task paradigm: The effects of providing visual speech cues. *Journal of Speech, Language and Hearing Research*, 53, 18-33.
- Fries, E., Dettenborn, L. & Kirschbaum, C. (2009). The cortisol awakening response (CAR): Facts and future directions. *International Journal of Psychophysiology*, 72, 67-73.
- Friesen, L. M., Shannon, R. V., Başkent, D., and Wang, X. (2001). Speech recognition in noise as a function of the number of spectral channels: Comparison of acoustic hearing and cochlear implants. *Journal of the Acoustical Society of America*, 110, 1150–1163.
- Fu, Q.-J. & Shannon, R. V. (1999). Recognition of spectrally degraded and frequency-shifted vowels in acoustic and electric hearing. *Journal of the Acoustical Society of America*, 105, 1889-1900.
- Fu, Q. -J. & Nogaki, G. (2005). Noise susceptibility of cochlear implant users: the role of spectral resolution and smearing. *Journal of the Association for Research in Otolaryngology*, 6, 19-27.
- Fu, Q.-J. & Galvin, J. J. (2008). Maximising cochlear implant patients' performance with advanced speech training procedures. *Hearing Research*, 242, 198-208.
- Gallacher, J., Ilubaera, V., Ben-Shlomo, Y., Bayer, A., Fish, M., Babisch, W. & Elwood, P. (2012). Auditory threshold, phonologic demand, and incident dementia. *Neurology*, 79, 1583-1590.

- Galvan, A., Hare, T. A., Parra, C. E., Penn, J., Voss, H., Glover, G. & Casey, B. J (2006). Earlier development of the accumbens relative to the orbitofrontal cortex might underlie risk-tasking behaviour in adolescents. *Journal of Neuroscience*, 26, 6885-6892.
- Ganesh, A. C., Vijitha, S. & Subba, R. T. A. (2011) Objective measurement of listening effort while using first and second language in simulated cochlear implants. Proceedings of ISAAR 2011: *Speech perception and auditory disorders*, pp. 127-134.
- Garadat, S. H., Zwolan, T. A. & Pfingst, B. E. (2013). Using temporal modulation sensitivity to select stimulation sites for processor MAPs in cochlear implant users. *Audiology & Neurotology*, 18, 247-260.
- Garnham, C., O'Driscoll, M., Ramsden A. R. & Saeed, S. (2002). Speech understanding in noise with a Med-El COMBI 40+ cochlear implant using reduced channel sets. *Ear & Hearing*, 23, 540-542.
- Gatehouse, S. & Gordon, J. (1990). Response times to speech stimuli as measures of benefit from amplification. *British Journal of Audiology*, 24, 63-68.
- Gatehouse, S. & Noble, W. (2004). The Speech, Spatial and Qualities of Hearing Scale (SSQ). *International Journal of Audiology*, 43, 85-99.
- Geers, A. & Brenner, C. (2003). Background and educational characteristics of prelingually deaf children implanted by five years of age. *Ear & Hearing*, 24, 2S-14S.
- Geers, A., Brenner, C. & Tobey, E. A. (2011a). Long-term outcomes of cochlear implantation in early childhood: Sample characteristics and data collection methods. *Ear & Hearing*, 32, 2S-12S.
- Geers, A. E., Strube, M. J., Tobey, E. A., Pisoni, D. B., Moog, J. S. (2011b). Epilogue: factors contributing to long-term outcomes of cochlear implantation in early childhood. *Ear & Hearing*, 32, 84S-92S.
- Gifford, R. H., Olund, A. P. & Dejong, M. (2011). Improving speech perception in noise for children with cochlear implants. *Journal of the American Academy of Audiology*, 22, 623-632.
- Goosens, K. A. & Maren, S. (2002). Long-term potentiation as a substrate for memory: Evidence from studies of amygdaloid plasticity and Pavlovian fear conditioning. *Hippocampus*, 12, 592-599.
- Gordon, K. A., Chaikof, M. H., Salloum, C., Goulding, G. & Papsin, B. (2012). Toward a method for programming balanced bilateral cochlear implant stimulation levels in children. *Cochlear Implants International*, 13, 220-227.

- Gosselin, P. A. & Gagné, J. –P. (2010). Use of a dual-task paradigm to measure listening effort. *Canadian Journal of Speech Language Pathology and Audiology*, 34, 43-51.
- Gosselin, P. A. & Gagné, J. –P. (2011a). Older adults expend more listening effort than young adults recognising speech in noise. *Journal of Speech, Language and Hearing Research*, 54, 944-958.
- Gosselin, P. A. & Gagné, J. –P. (2011b). Older adults expend more listening effort than young adults recognising audiovisual speech in noise. *International Journal of Audiology*, 50, 786-792.
- Goupell, M. J., Stoelb, C., Kan, A. & Litovsky, R. Y. (2013). Effect of mismatched place-of-stimulation on the salience of binaural cues in conditions that simulate bilateral cochlear-implant listening. *Journal of the Acoustical Society of America*, 133, 2272-2287.
- Goy, H., Pelletier, M., Coletta, M. & Pichora-Fuller, M. K. (2013). The effects of semantic context and the type and amount of acoustic distortion on lexical decision by younger and older adults. *Journal of Speech, Language and Hearing Research*, 56, 1715-1732.
- Grady, C. (2012). The cognitive neuroscience of ageing. *Nature Reviews Neuroscience*, 13, 441-505.
- Granholm, E., Asarnow, R. F., Sarkin, A. J. & Dykes, K. L. (1996). Pupillary responses index cognitive resource limitations. *Psychophysiology*, 33, 457-461.
- Green, D. M. & Swets, J.A. (1966). *Signal Detection Theory and psychophysics*. New York, NY: Wiley.
- Green, K. M. J., Bhatt, Y. M., Mawman, D. J., O'Driscoll, M. P., Saeed, S. R., Ramsden, R. T. & Green, M. W. (2007). Predictors of audiological outcome following cochlear implantation in adults. *Cochlear Implants International*, 8, 1-11.
- Greenwood, D. D. (1990). A cochlear frequency-position function for several species—29 years later. *Journal of the Acoustical Society of America*, 87, 2592–2605.
- Grimby, A. & Ringdahl, A. (2000). Does having a job improve the quality of life among post-lingually deafened Swedish adults with severe-profound hearing impairment. *British Journal of Audiology*, 34, 187-195.
- Grossi, G., Perski, A., Ekstedt, M., Johansson, T., Lindström, M. & Holm, K. (2005). The morning salivary cortisol response to burnout. *Journal of Psychosomatic Research*, 59, 103-111.
- Guinan, J. J. Jr. (1996). Physiology of olivocochlear efferents. In P. Dallos, A. N. Popper & R. R. Fay (Eds.), *The cochlea*. (pp. 435-502). New York, NY: Springer.

- Hällgren, M., Larsby, B., Lyxell, B. & Arlinger, S. (2001). Evaluation of a cognitive test battery in young and elderly normal-hearing and hearing-impaired persons. *Journal of the American Academy of Audiology*, 12, 357-370.
- Hammen, C. (2005). Stress and depression. *Annual Review of Clinical Psychology*, 1, 293-319.
- Hardy, M. Y. & Tye-Din, J. A. (2016). Coeliac disease: a unique model for investigating broken tolerance in autoimmunity. *Clinical & Translational Immunology*, 5, e112-e122.
- Hartmann, R., Topp, G. & Klinke, R. (1984). Discharge patterns of cat primary auditory nerve fibers with electrical stimulation of the cochlea. *Hearing Research*, 13, 47-62.
- Hartmann, R. & Kral, A. (2004). *Central responses to electrical stimulation*. In F. G. Zeng, A. N. Popper & R. R. Fay (Eds.), *Cochlear implants: auditory prostheses and electric hearing* (pp. 213-285). New York, NY: Springer.
- Hayes, T. R. & Petrov, A. A. (2015). Mapping and correcting the influence of gaze position on pupil size measurements. *Behaviour Research Methods*, 48, 510-527.
- Hebb, D. O. (1955). *Drives and the C.N.S. (conceptual nervous system)*. *Psychological Review*, 62, 243-254.
- Heeger, D. (1997). *Signal Detection Theory*. Retrieved from <http://www.cns.nyu.edu/~david/handouts/sdt-advanced.pdf>
- Heitz, R. P., Schrock, J. C., Payne, T. W. & Engle, R. W. (2008). Effects of incentive on working memory capacity: Behavioral and pupillometric data. *Psychophysiology*, 45, 119-129.
- Henry, J. P., Stephens, P. M., Santisteban, G. A. (1975). A model of psychosocial hypertension showing reversibility and progression of cardiovascular complications. *Circulation Research*, 36, 156-164.
- Hersbach, A.A., Grayden, D. B., Fallon, J. B. & McDermott, H, J. (2013). A beamformer post-filter for cochlear implant noise reduction. *The Journal of the Acoustical Society of America*, 133, 2412-2420.
- Hervais-Adelman, A., Davis, M. H., Johnsrude, I. S. & Carlyon, R. P. (2008). Perceptual learning of noise-vocoded words: Effects of feedback and lexicality. *Journal of Experimental Psychology: Human Perception and Performance*, 34, 460-474.
- Hervais-Adelman, A. G., Davis, M. H., Johnsrude, I. S., Taylor, K. J. & Carlyon, R. P. (2011). Generalization of perceptual learning of vocoded speech. *Journal of Experimental Psychology: Human Perception and Performance*, 37, 283-295.

- Hetú, R., Riverin, L., Lalande, N., Getty, L., & St-Cyr, C. (1988). Qualitative analysis of the handicap associated with occupational hearing loss. *British Journal of Audiology*, 22, 251-264.
- Hicks, C. B. & Tharpe, A. M. (2002). Listening effort and fatigue in school-age children with and without hearing loss. *Journal of Speech, Language and Hearing Research*, 45, 573-584.
- Hindley, P. (2000). *Child and adolescent psychiatry*. In P. Hindley & N. Kitson (Eds.), *Mental Health and Deafness*. London, UK: Whurr Publishers.
- Hindley, P., Hill, P., McGuigan, S. & Kitson, N. (1994). Psychiatric disorder in deaf and hearing impaired children & young people: a prevalence study. *Journal of Child Psychology and Psychiatry*, 55, 917-934.
- Hinojosa, R. & Marion, M. (1983). Histopathology of profound sensorineural deafness. *Annals of the New York Academy of Sciences*, 405, 459-484.
- Hochmair, I. (2013). *Cochlear Implants: Facts*. Retrieved from <http://www.medel.com/cochlear-implants-facts/>
- Hochmair, I., Nopp, P., Jolly, C., Schmidt, M., Schösser, H., Garnham, C. & Anderson, I. (2006). MED-EL Cochlear implants: State of the art and a glimpse into the future. *Trends in Amplification*, 10, 201-219.
- Hockey, R. (2013). *The Psychology of Fatigue: Work, Effort and Control*. Cambridge, UK: Cambridge University Press.
- Holden, L.K., Finley, C. C., Firszt, J. B., Holden, T. A., Brenner, C., Potts, L. G., Gotter, B. D., Vanderhoof, S. S., Mispagel, K. , Heydebrand, G. (2013). Factors affecting open-set word recognition in adults with cochlear implants. *Ear & Hearing*, 34, 342-360.
- Holsboer, F. (2000). The corticosteroid receptor hypothesis of depression. *Neuropsychopharmacology*, 23, 477-501.
- Hopstaken, J. F., van der Linden, D., Bakker, A. B. & Kompier, M. A. J. (2015). The window of my eyes: Task disengagement and mental fatigue covary with pupil dynamics. *Biological Psychology*, 110, 100-106.
- Hornsby, B. W. (2013). The effects of hearing-aid use on listening effort and mental fatigue associated with sustained speech processing demands. *Ear & Hearing*, 34, 523-534.
- Hornsby, B. W. Y., Werfel, K., Camarta, S. & Bess, F. H. (2014). Subjective fatigue in children with hearing loss: some preliminary findings. *American Journal of Audiology*, 23, 129-134.
- Hornsby, B. W. Y., Naylor, G. & Bess, F. H. (2016). A taxonomy of fatigue concepts and their relation to hearing loss. *Ear & Hearing*, 37, 136S-144S.

- Houben, R., Van Doorn-Bierman, M. & Dreschler, W. A. (2013). Using response time to speech as a measure for listening effort. *International Journal of Audiology*, 52, 753-761.
- Howard, C. S., Munro, K. J., & Plack, C. J. (2010). Listening effort at signal-to-noise ratios that are typical of the school classroom. *International Journal of Audiology*, 49, 928–932.
- Hübener, M. & Bonhoeffer, T. (2014). Neuronal plasticity: beyond the critical period. *Cell*, 159, 727-737.
- Huckvale, M. (2013). *SFS/Window: Speech Filing System (Version 4.10)*. Available from <http://www.phon.ucl.ac.uk/resource/sfs/>
- Hughes, K. C. & Galvin, K. L. (2013). Measuring listening effort expended by adolescents and young adults with unilateral or bilateral cochlear implants or normal hearing. *Cochlear Implants International*, 14, 121-129.
- Hughes, S. E., Rapport, F. L., Boisvert, I., McMahon, C. M. & Hutchings, H. A. (2017). Patient-reported outcome measures (PROMs) for assessing perceived listening effort in hearing loss: protocol for a systematic review. *BMJ Open*, 0, 1-5.
- Humes, L. E. (1999). Dimensions of hearing aid outcome. *Journal of the American Academy of Audiology*, 10, 26-39.
- Humes, L. E. & Bess, F. H. (2013). *Audiology and Communication Disorders: An overview*. (2nd ed.) Baltimore, MD: Lippincott, Williams & Wilkins.
- Humes, L. & Humes, L (2004). Factors affecting long-term hearing aid success. *Seminars in Hearing*, 25, 63-72.
- Intraub, H. (1980). Presentation rate and the representation of briefly glimpsed pictures in memory. *Journal of Experimental Psychology: Human Perception & Performance*, 6, 1-12.
- Jahncke, H. & Halin, N. (2012). Performance, fatigue and stress in open-plan offices: The effects of noise and restoration on hearing impaired and normal hearing individuals. *Noise Health*, 14, 260-272.
- Janisse, M. P. (1977). *Pupillometry: The Psychology of the Pupillary Response*. Washington, D.C: Hemisphere Publishing Co.
- Järvelin, M. R., Mäki-Torodo, E., Sorri, M. J. & Rantakallio, P. T. (1997). Effect of hearing impairment on educational outcomes and employment up to the age of 25 years in northern Finland. *British Journal of Audiology*, 31, 165-175.
- Jerjes, W. K., Cleare, A. J., Wessely, S., Wood, P. J. & Taylor, N. F. (2005). Diurnal patterns of salivary cortisol and cortisone output in chronic fatigue syndrome. *Journal of Affective Disorders*, 87, 299-304.
- Kaernbach, C. (1991). Simple adaptive testing with the weighted up-down method. *Perception & Psychophysics*, 49, 227-229.

- Kahneman, D. (1973). *Attention and effort*. Englewood Cliffs, NJ: Prentice-Hall.
- Kahneman, D. (2011). *Thinking, fast and slow*. New York, NY: Farrar, Straus and Giroux.
- Kahneman, D., Oonuska, L. & Wolman, R. E. (1968). Effects of grouping on the pupillary response in a short-term memory task. *Quarterly Journal of Experimental Psychology*, 20, 309-3011.
- Karatekin, C. (2004). Development of attentional allocation in the dual-task paradigm. *International Journal of Psychophysiology*, 52, 7-21.
- Kato, T. A., Hayakawa, K., Monji, A. & Kanba, S. (2013). Missing and possible link between neuroendocrine factors, neuropsychiatric disorders and microglia. *Frontiers in Integrative Neuroscience*, 7, 1-16.
- Kerlin, J. R., Shahin, A. J. & Miller, L. M. (2010). Attentional gain control of on-going cortical speech representations in a “cocktail party”. *Journal of Neuroscience*, 30, 620-628.
- Khan, A. M., Handzel, O., Burgess, B. J., Damian, D., Eddington, D. K. & Nadol, J. B., Jr. (2005). Is word recognition correlated with the number of surviving spiral ganglion cells and electrode insertion depth in human subjects with cochlear implants? *Laryngoscope*, 115, 672-677.
- Kiang, N. Y., Watanabe, T., Thomas, E. C. & Clark, L. F. (1965). *Discharge patterns of single fibres in the cat's auditory nerve*. Cambridge, MA: MIT Press.
- Kiang, N. Y., Moxon, E. C. & Levine, R. A. (1970). Auditory-nerve activity in cats with normal and abnormal cochleas. In G. E. Wolstenholme & J. Knight (Eds.), *Sensorinueral hearing loss* (pp. 241-273). London, UK: Churchill.
- Kiang, N. Y., Guinan, J. J. Jr., Liberman, M. C., Brown, M. C. & Eddington, D. K. (1988). Feedback control mechanisms for the auditory periphery: Implications for cochlear implants. In P. Banfai, P. (Ed.), *Cochlear implant: Current situation* (pp. 131-141). Erkelenz, Germany: Bermann.
- Kiecolt-Glaser, J. K., McGuire, L., Robles, T. F. & Glaser, R. (2002). Psychoneuroimmunology: psychological influences on immune function and health. *Journal of Consulting and Clinical Psychology*, 70, 537-547.
- Kiefer, J., Von Ilberg, C., Rupprecht, V., Hubner-Egener, J. & Knecht, R. (2000). Optimized speech understanding with the continuous interleaved sampling speech coding strategy in patients with cochlear implants: Effects of variations in stimulation rate and number of channels. *Annals of Otology, Rhinology & Laryngology*, 109, 1009-1020.
- Kinsbourne, M. & Warrington, E. K. (1962). The effect of an aftercoming random pattern on the perception of brief visual stimuli. *Quarterly Journal of Experimental Psychology*, 14, 223-234.

- Klein, S. A. (2001). Measuring, estimating and understanding the psychometric function: A commentary. *Perception and Psychophysics*, 68, 1421-1455.
- Koch, D. B., Downing, M., Osberger, M. J., Litvak, L. (2007). Using current steering to increase spectral resolution in CII and HiRes 90K users. *Ear & Hearing*, 28, 39S-41S.
- Koelewijn, T., Zekveld, A. a., Festen, J. M., & Kramer, S. E. (2012). Pupil dilation uncovers extra listening effort in the presence of a single-talker masker. *Ear & Hearing*, 33, 291–300.
- Kouneiher, F., Charron, S. & Koechlin, E. (2009). Motivation and cognitive control in the human prefrontal cortex. *Nature Neuroscience*, 12, 939-945.
- Kral, A. (2013). Auditory critical periods: A review from system's perspective. *Neuroscience*, 247, 117-133.
- Kral, A. & O'Donoghue, G. M. (2010). Profound deafness in childhood. *New England Journal of Medicine*, 363, 1438-1450.
- Kral, A. & Sharma, A. (2012). Developmental neuroplasticity after cochlear implantation. *Trends in Neuroscience*, 35, 111-122.
- Kral, A., Hartmann, R., Tillen, J., Heid, S. & Klinke, R. (2002). Hearing after congenital deafness: central auditory plasticity and sensory deprivation. *Cerebral Cortex*, 12, 797-807.
- Kral, A., Tillein, J., Heid, S., Hartmann, R. & Klinke, R. (2005). Postnatal cortical development in congenital auditory deprivation. *Cerebral Cortex*, 15, 552-562.
- Kral., A., Kronenberger, W. G., Pisoni, D. B. & O'Donoghue, G. M. (2016). Neurocognitive factors in sensory restoration of early deafness: a connectome model. *The Lancet*, 15, 610-621.
- Kramer, S.E., Kapteyn, T.S. Festen, J. M. & Kuik, D. J. (1997). Assessing aspects of auditory handicap by means of pupil dilation. *Audiology*, 36, 155-164.
- Kramer, S.E., Kapteyn, T.S., Houtgast, T. (2006). Occupational performance: Comparing normally-hearing and hearing-impaired employees using the Amsterdam checklist for hearing and work. *International Journal of Audiology*, 45, 5003-5512.
- Kramer, S. E., Lorens, A., Coninx, F., Zekveld, A.A., Piotrowska, A. & Skarzynski, H. (2012). Processing load during listening: the influence of task characteristic on the pupil response. *Language and Cognitive Processes*, 28, 426-442.
- Kramer, S. E., Zekveld, A. & Koelewijn, T. (2015, August). *The First Summer Workshop on the Application of Pupillometry in Hearing Science to assess Listening Effort*. Workshop conducted from VU University Medical Center, Amsterdam, Netherlands.

- Kramer, S. E., Teunissen, C. E. & Zekveld, A. A. (2016). Cortisol, chromogranin A, and pupillary responses evoked by speech recognition tasks in normally hearing and hard-of-hearing listeners: a pilot study. *Ear & Hearing*, 37, 126S-135S.
- Kronenberger, W. G., Pisoni, D. B., Henning, S. C. & Colson, B. G. (2013). Executive functioning skills in long-term users of cochlear implants: a case control study. *Journal of Pediatric Psychology*, 38, 902-914.
- Kronenberger, W. G., Beer, J., Castellanos, I., Pisoni, D. B., Miyamoto, R. T. (2014). Neurocognitive risk in children with cochlear implants. *Journal of the American Medical Association: Otolaryngology-Head & Neck Surgery*, 140, 608-615.
- Kuchinsky, S. E., Ahlstrom, J. B., Vaden, I. I., Cute, S. L., Humes, L. E., Dubno, J.R. & Eckert, M. A. (2013). Pupil size varies with word listening and response selection difficulty in older adults with hearing loss. *Language and Cognitive Processing*, 28, 426-442.
- Kudielka, B. M., Bellingrath, S. & Hellhammer, D. H. (2006). Cortisol in burnout and vital exhaustion: An overview. *Giornale Italiano Di Medicina Del Lavoro Ed Ergonomia*, 28, 34-42.
- Kudielka, B. M., Helhammer, D. H. & Wüst, S. (2009). Why do we respond so differently? Reviewing determinants of human salivary cortisol responses to challenge. *Psychoneuroendocrinology*, 34, 2-18.
- Kumari, M., Badrick, E., Chandola, T., Adam, E. K., Stafford, M., Marmot, M. G., Kirschbaum, C. & Kivimaki, M. (2009). Cortisol secretion and fatigue: Associations in a community based cohort. *Psychoneuroendocrinology*, 34, 1476-1485.
- Kunugi, H., Ida, I., Owashi, T., Kimura, M., Inoue, Y., Nakagawa, S., Yabana, T., Urushibara, T., Kanai, R., Aihara, M., Yuuki, N., Otsubo, T., Oshima, A., Kudo, K., Inoue, T., Kitaichi, Y., Shirakawa, O., Isogawa, K., Nagayama, H., Kamijima, K., Nanko, S., Kanba, S., Higuchi, T. & Mikuni, M. (2006). Assessment of the dexamethasone/CRH test as a state-dependent marker for hypothalamic-pituitary-adrenal (HPA) axis abnormalities in major depressive episode: a Multicenter Study. *Neuropsychopharmacology*, 31, 212-220.
- Laeng, B., Sirios, S. & Gredebäck, G. (2012). Pupillometry: A window to the preconscious? *Perspectives in Psychological Sciences*, 7, 18-27.
- Lamb, B. & Archbold, S. (2013). *Adult cochlear implantation: The case for a review of provision*. Ear Foundation, 1-20.
- Lane, M. & Mackersie, C. L. (2015). Sound tolerance domains: Effects of noise type. Proceedings from: *American Auditory Society*. Scottsdale, AZ.

- Larsby, B., Hällgren, M., Lyxell, B. & Arlinger, S. (2005). Cognitive performance and perceived effort in speech processing tasks: Effects of different noise backgrounds in normal-hearing and hearing-impaired subjects. *International Journal of Audiology*, 44, 131-143.
- Lawson, D. T., Wilson, B. S., Zerbi, M. & Finley, C. C. (1996). Speech processors for auditory prostheses. *Third Quarterly Progress Report, NIH project N01-DC-5-2103*. Bethesda, MD: Neural Prosthesis Program, National Institutes of Health.
- Lazard, D. S., Bordure, P., Lina-Granade, G., Magnam, J., Meller, B., Radafy, E., Roux, P. E., Gnansia, D., Pean, V., Truy, E. (2010). Speech perception performance for 100 post-lingually deaf adults fitted with NeureJec cochlear implants: comparison between Digisonic® Convex and Digisonic® SP devices after a 1-year follow-up. *Acta Oto-laryngologica*, 130, 1267-1273.
- Lazard, D. S., Vincent, C., Venail, F., Van de Heyning, P., Truy, E., Sterkers, O., Skarzynski, P. H., Skarzynski, H., Schauwers, K., O’Leary, S., Mawman, D., Maat, B., Kleine-Punte, A., Huber, A. M., Green, K., Govaerts, P.J., Fraysse, B., Dowell, R., Dillier, N., Burke, E., Beynon, A., Bergeron, F., Başkent, D., Artieries, F. & Blamey, P. J. (2012a). Pre-, per- and postoperative factors affecting performance of post-linguistically deaf adults using cochlear implants: a new conceptual model over time. *PLoS One*, 7, 1-11.
- Lazard, D. S., Giraud, A.-L., Gnansia, D., Meyer, B. & Sterkers, O. (2012b). Understanding the deafened brain: Implications for cochlear implant rehabilitation. *European Annals of Otorhinolaryngology, Head and Neck diseases*, 129, 98-103.
- Lazarus, R. S. & Folkman, S. (1984). *Stress, appraisal and coping*. New York, NY: Springer.
- Leake, P. A. & Rebscher, S. J. (2004). Anatomical considerations and long-term effects of electrical stimulation. In F. G. Zeng, A. N. Popper & R. R. Fay (Eds.), *Auditory prostheses: Cochlear implants and beyond* (pp. 101-148). New York, NY: Springer-Verlag.
- Lee, H. J., Kang, E., Oh, S. –H., Kang, H., Lee, D. S., Lee, M. C. & Kim, C. –S. (2005). Preoperative differences of cerebral metabolism relate to the outcome of cochlear implants in congenitally deaf children. *Hearing Research*, 203, 2-9.
- Lee, R. S., Tamashiro, K. L., Yang, X., Purcell, R. H., Harvey, A., Willour, V. L., Huo, Y., Rongione, M., Wand, G. S. & Potash, J. B. (2010). Chronic corticosterone exposure increases expression and decreases deoxyribonucleic acid methylation of Fkbp5 in mice. *Endocrinology*, 151, 4332-4343.
- Leek, M. R. (2001). Adaptive procedures in psychophysical research. *Perception & Psychophysics*, 63, 1279-1292.

- Lemke, U. & Besser, J. (2016). Cognitive load and listening effort: Concepts and age-related considerations. *Ear & Hearing*, 37, 77S-84S.
- Lesch, K. P., Mayer, S., Disselkamp-Tietze, J., Hoh, A., Wiesmann, M., Osterheider, M. & Schulte, H. M. (1990). 5-HT_{1A} receptor responsivity in unipolar depression. Evaluation of ipsapirone-induced ACTH and cortisol secretion in patients and controls. *Biological Psychiatry*, 28, 620-628.
- Levitt, H. (1971). Transformed up-down methods in psychoacoustics. *The Journal of the Acoustical Society of America*, 49, 467-477.
- Lewis, R. L., Vasishth, S. & Van Dyke, J. A. (2006). Computational principles of working memory in sentence comprehension. *Trends in Cognitive Sciences*, 10, 447-454.
- Liao, H. -I., Kidani, S., Yoneya, M., Kashino, M. & Furukawa, S. (2015). Correspondences among pupillary dilation response, subjective salience of sounds, and loudness. *Psychonomic Bulletin & Review*, 23, 412-425.
- Liberman, M. C. (1978). Auditory-nerve response from cats raised in a low-noise chamber. *Journal of the Acoustical Society of America*, 63, 442-455.
- Lin, F. R. & Albert, M. (2014). Hearing loss and dementia – who is listening? *Aging & Mental Health*, 18, 671-673.
- Lin, F. R., Metter, E. J., O'Brien, R. J., Resnick, S. M., Zonderman, A. B. & Ferrucci, L. (2011). Hearing loss and incident dementia. *Archives of Neurology*, 68, 214-220.
- Lin, F. R., Yaffe, K., Xia, J., Xue, Q. -L., Harris, T. B., Purchase-Helzner, E., Satterfield, S., Ayonayon, H. N., Ferrucci, L. & Simonsick, E. M. (2013). Hearing loss and cognitive decline among older adults. *Journal of the American Medical Association Internal Medicine*, 173, 293-299.
- Liston, C., Miller, M. M., Goldwater, D. S., Radley, J. J., Rocher, A. B., Hof, P. R., Morrison, J. H. & McEwen, B. S. (2006). Stress-induced alterations in prefrontal cortical dendritic morphology predict selective impairments in perceptual attentional set-shifting. *Journal of Neuroscience*, 26, 7870-7874.
- Litovsky, R. Y., Parkinson, A. & Arcaroli, J. (2009). Spatial hearing and speech intelligibility in bilateral cochlear implant users. *Ear & Hearing*, 30, 419-431.
- Litovsky, R. Y., Goupell, M. J., Godar, S., Grieco-Calub, T., Jones, G. L., Garadat, S. N., Agrawal, S., Kan, A., Todd, A., Hess, C. & Misurelli, S. (2012). Studies on bilateral cochlear implants at the University of Wisconsin's Binaural Hearing and Speech Laboratory. *Journal of the American Academy of Audiology*, 23, 476-494.
- Loewenfeld, I. E. (1993). *The pupil: anatomy, physiology, and clinical application*, Volume I. Ames, IA: Iowa State UP.

- Loizou, P. C., Dorman, M. & Tu, Z. (1999). On the number of channels needed to understand speech. *Journal of the Acoustical Society of America*, 106, 2097-2103.
- Loizou, P. C., Lobo, A. & Hu, Y. (2005). Subspace algorithms for noise reduction in cochlear implants (L). *Journal of the Acoustical Society of America*, 118, 2791-2793.
- Luce, R. D. (1963). A threshold theory for simple detection experiments. *Psychological Review*, 70, 61-79.
- Luce, R. D. (1993). *Sound and hearing: A conceptual introduction*. Hillsdale, NJ: Erlbau.
- Luks, T. L., Simpson, G. V., Dale, C. L. & Hough, M. G. (2007). Preparatory allocation of attention and adjustments in conflict processing. *Neuroimage*, 35, 949-958.
- Lunner, T. (2003). Cognitive function in relation to hearing-aid use. *International Journal of Audiology*, 42, 49-58.
- Lunner, T., Rudner, M., Rosenbom, T., Ågren, J. & Ng, E. H. N. (2016). Using speech recall in hearing aid fitting and outcome evaluation under ecological test conditions. *Ear & Hearing*, 37, 145S-154S.
- Lupien, S., Sindi, S. & Wan, N. (2012). *When we test, do we stress? Guidelines for health professionals and scientists working with older adults*. Montreal, Quebec: Centre for Studies on Human Stress. Retrieved from http://www.humanstress.ca/documents/pdf/KT/KT_document_EN.pdf.
- Lurie, A. & Lurie, R. H. (2014). *Data logging in pediatric cochlear implant recipients*. Retrieved from https://cdn.ymaws.com/www.acialliance.org/resource/resmgr/CI2014_PPT/tour_nis_fri.pdf.
- Luts, H., Eneman, K., Wouters, J., Schulte, M., Vormann, M. & Spriet, A.(2010). Multicenter evaluation of signal enhancement algorithm for hearing aids. *Journal of the Acoustical Society of America*, 127, 1491-1505.
- Lyness, C.R., Woll, B., Campbell, R. & Cardin, V. (2013). How does visual language affect crossmodal plasticity and cochlear implant success? *Neuroscience and Biobehavioural Reviews*, 37, 2621-2630.
- Lyxell, B., Sahlén, B., Wass, M., Ibertsson, T., Larsby, B., Hällgren, M. & Mäki-Torkko, E. (2008). Cognitive development in children with cochlear implants: relations to reading and communication. *International Journal of Audiology*, 47 (supp 2), 47S-52S.
- Machin, D., Campbell, M., Fayers, P. & Pinol, A. (1997). *Sample size tables for clinical studies* (2nd ed.) Malden, MA: Blackwell Science.

- Mackersie, C. L. & Cones, H. (2011). Subjective and psychophysiological indexes of listening effort in a competing-talker task. *Journal of the American Academy of Audiology*, 22, 113-122.
- Mackersie, C. L., Baxter, S. & Lane, M. (2014). Domains of sound-tolerance underlying Acceptable Noise Level. Proceedings from IHCON '14: *International Hearing Aid Research Conference*. Lake Tahoe, CA.
- Mackersie, C. L. & Calderon-Moultrie, N. (2016). Autonomic nervous system reactivity during speech repetition tasks: heart rate variability and skin conductance. *Ear & Hearing*, 37, 118S-125S.
- MacPherson, A. & Akeryod, M. A. (2013). The Glasgow monitoring of uninterrupted speech task: A naturalistic measure of speech intelligibility in noise. *Journal of the Acoustical Society of America*. 133, 3379.
- Magarino, A. M. & McEwen, B. S. (1995). Stress-induced atrophy of apical dendrites of hippocampal CA3c neurons: involvement of glucocorticoid secretion and excitatory amino acid receptors. *Neuroscience*, 69, 89-98.
- Marinelli, M. & Piazza, P. V. (2002). Interaction between glucocorticoid hormones, stress and psychostimulant drugs. *European Journal of Neuroscience*, 16, 387-394.
- Matthen, M. (2016). Effort and displeasure in people who are hard of hearing. *Ear & Hearing*, 37, 28S-34S.
- Mattys, S. L., Brooks, J. & Cooke, M. (1990). Recognizing speech under a processing load: Dissociating energetic from informational factors. *Cognitive Psychology*, 59, 203-243.
- Mattys, S. L. & Wiget, L. (2011). Effect of cognitive load on speech recognition. *Journal of Memory and Language*, 65, 145-160.
- Mattys, S. L., Davis, M. H., Bradlow, A. R. & Scott, S. K. (2012). Speech recognition in adverse conditions: A review. *Language and Cognitive Processes*, 27, 953-978.
- Mauger, S. J., Arora, K. & Dawson, P. W. (2012). Cochlear implant optimized noise reduction. *Journal of Neural Engineering*, 9.
- McBride, E. G., Rubel, E. W. & Wang, Y. (2013). Afferent regulation of chicken auditory brainstem neurons: rapid changes in phosphorylation of elongation factor 2. *Journal of Comparative Neurology*, 521, 1165-1183.
- McCoy, S. L., Tun, P. A., Cox, L. C., Colangelo, M., Stewart, R. A. & Wingfield, A. (2005). Hearing loss and perceptual effort: Downstream effects on older adults' memory for speech. *Quarterly Journal for Experimental Psychology*, 58A, 22-33.
- McEwen, B. S. (1998). Protective and damaging effects of stress mediators. *New England Journal of Medicine*, 338, 171-179.

- McEwen, B. S. & Gianaros, P. J. (2010). Central role of the brain in stress and adaptation: Links to socioeconomic status, health and disease. *Annals of the New York Academy of Sciences*, 1186, 190-222.
- McFadden, B. & Pittman, A. (2008). Effect of minimal hearing loss on children's ability to multitask in quiet and in noise. *Language, Speech and Hearing Services in Schools*, 39, 342-351.
- McGarrigle, R., Munro, K. J., Dawes, P., Stewart, A. J., Moore, D. R., Barry, J. G. & Amitay, S. (2014). Listening effort and fatigue: What exactly are we measuring? A British Society of Audiology Cognition in Hearing Special Interest Group 'white paper'. *International Journal of Audiology*, 53, 433-445.
- McKlveen, J. M., Myers, B. & Herman, J. P. (2015). The medial prefrontal cortex: coordinator of autonomic, neuroendocrine, and behavioural responses to stress. *Journal of Neuroendocrinology*, 27, 446-456.
- Menon, V. & Uddin, L. Q. (2010). Saliency, switching, attention and control: A network model of insula function. *Brain Structure and Function*, 214, 655-667.
- Miller, G. A. (1947). The masking of speech, *Psychological Bulletin*, 44, 105-129.
- Miller, E. K. & Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Annual Review of Neuroscience*, 24, 167-202.
- Miura, M., Sando, I., Hirsch, B. E. & Orita, Y. (2002). Analysis of spiral ganglion cell populations in children with normal and pathological ears. *Annals of Otolaryngology, Rhinology & Laryngology*, 111, 1059-1065.
- Monk, C. S., McClure, E. B., Nelson, E. E., Zarahn, E., Bilder, R. M., Leibenluft, E., Charney, D. S., Ernst, M. & Pine, D. S. (2003). Adolescent immaturity in attention-related brain engagement to emotional facial expressions. *Neuroimage*, 20, 420-428.
- Moore, D.R. & Shannon, R. V. (2009). Beyond cochlear implants: Awakening the deafened brain. *Nature Neuroscience*, 12, 686-691.
- Morgan, J. T., Chana, G., Pardo, C. A., Achim, C., Semendeferi, K., Buckwalter, J., Courchesne, E. & Everall, I. P. (2010). Microglial activation and increased microglial density observed in the dorsolateral prefrontal cortex in autism. *Biological Psychiatry*, 68, 368-376.
- Morgan, J. T., Chana, G., Abramson, I., Semendeferi, K., Courchesne, E. & Everall, I. P. (2012). Abnormal microglial-neuronal spatial organisation in the dorsolateral prefrontal cortex in autism. *Brain Research*, 1456, 72-81.
- Munson, B., Donaldson, G. S., Allen, S. L., Collison, E. A. & Nelson, D. A. (2003). Patterns of phoneme misperceptions by individuals with cochlear implants. *Journal of the Acoustical Society of America*, 113, 925-935.

- Murphy, P. R., Robertson, I. H., Balsters, J. H. & O'Connell, R. G. (2011). Pupillometry and P3 index the locus coeruleus-noradrenergic arousal function in humans. *Psychophysiology*, 48, 1532-1543.
- Nachtegaal, J., Kuik, D. J., Anema, J.R., Goverts, S. T., Festen, J. M. & Kramer, S. E. (2009). Hearing status, need for recovery after work, and psychosocial work characteristics: Results from an internet-based national survey on hearing. *International Journal of Audiology*, 48, 684-491.
- Nadol, J. B. Jr., Shiao, J. Y., Burgess, B. J., Ketten, D. R., Eddington, D. K., Gantz, B. J., Kos, I., Montandon, P., Coker, N. J., Roland, J. T. Jr. & Shallop, J. K. (2001). Histopathology of cochlear implants in humans. *Annals of Otology, Rhinology & Laryngology*, 110, 883-891.
- NDCS (2016). How do I help my child cope with tiredness? *NDCS Families Winter Issue*. pp.28-29.
- NDCS (2017). *How radio aids can help: A guide for families*. pp.1-40.
- NDCS (2018). *Getting your child to wear their hearing aids and cochlear implant processors*. Retrieved from http://www.ndcs.org.uk/family_support/childhood_deafness/hearing_aids/getting_your_child.html.
- Nelson, P. B., Jin, S-H., Carney, A. E., & Nelson, D. A. (2003). Understanding speech in modulated interference: Cochlear implant users and normal-hearing listeners. *Journal of the Acoustical Society of America*, 113, 961–968.
- Nie, K., Stickney, G. & Zeng, F. G. (2005). Encoding frequency modulation to improve cochlear implant performance in noise. *IEEE Transactions on Biomedical Engineering*, 52, 64-73.
- NIH Blueprint (2010). *\$40 million awarded to trace human brain connections*. Retrieved from <https://www.nih.gov/news-events/news-releases/40-million-awarded-trace-human-brains-connections>
- Nijhof, S. L., Rutten, J. M. T. M., Uiterwaal, C. S. P. M., Bleijenberg, G., Kimpen, J. L. L. & van de Putte, E. M. (2014). The role of hypocortisolism in chronic fatigue syndrome. *Psychoneuroendocrinology*, 42, 199-206.
- Niparko, J. K., Tobey, E.A., Thal, D. J., Eisenberg, L.S., Wang, N.-Y., Quittner, A. L., Fink, N. E. (2010). Spoken language development in children following cochlear implantation. *Journal of the American Medical Association*, 303, 1498-1506.
- Nittrouer, S., Caldwell-Tarr, A. & Lowenstein, J. H. (2013). Working memory in children with cochlear implants: problems are in storage, not processing. *International Journal of Pediatric Otorhinolaryngology*, 77, 1886-1898.

- Niwa, M., Jaaro-Peled, H., Tankou, S., Seshadri, S., Hikida, T., Matsumoto, Y., Cascella, N. G., Kano, S. –I., Ozaki, N., Nabeshima, T. & Sawa, A. (2013). Adolescent stress-induced epigenetic control of dopaminergic neurons via glucocorticoids. *Science*, 339, 335-339.
- Noble, W. & Gatehouse, S. (2006). Effects of bilateral versus unilateral hearing aid fitting on abilities measured by the Speech, Spatial and Qualities of Hearing Scale (SSQ). *International Journal of Audiology*, 43, 172-181.
- Nusbaum, H. & Morin, T. (1992). Paying attention to differences among talkers. In Y. Tohkura, E. Bateson & Y. Sagisaka (Eds). *Speech perception, production and linguistic structure*. Pp 66-94. Tokyo: IOS Press.
- Obleser, J. & Kotz, S. A. (2011). Multiple brain signatures of integration in the comprehension of degraded speech, *NeuroImage*, 55, 713-723.
- Obleser, J., Wöstmann, M., Helbernd, N., Wilsch, A. & Maess, B. (2012). Adverse listening conditions and memory load drive a common alpha oscillatory network. *Journal of Neuroscience*, 32, 12376-12383.
- O'Donoghue, G. M., Nikolopoulos, T. P. & Archbold, S. M. (2000). Determinants of speech perception in children after cochlear implantation. *Lancet*, 356, 466-468.
- O'Donoghue, G., Kral, A. J., Pisoni, D. & Kronenberger, W. (2016). The brain's connectome – a symphony inside our brains and how hearing loss disturbs the music. *ENT and Audiology News*, 25, 2-4.
- Okusa, M., Shiraishi, T., Kubo, T., Nageishi, Y. (1999). Effects of discrimination difficulty on cognitive event-related potentials in patients with cochlear implants. *Otolaryngology, Head and Neck Surgery*, 121, 610-615.
- Ophir, E., Nass, C. I. & Wagner, A. D. (2009). Cognitive control in media multitaskers. *Proceedings of the National Academy of Sciences*, 106, 15583-15587.
- Paas, F., Tuovinen, J. E., Tabber, H. & Van Gerven, P. W. M. (2003). Cognitive load measurement as a means to advance cognitive load theory. *Educational Psychologist*, 38, 63-71.
- Pals, C., Sarampalis, A. & Başkent, B. (2013). Listening effort with cochlear implant simulations. *Journal of Speech, Language and Hearing Research*, 56, 1075-1084.
- Parker, A. J., Wessely, S. & Cleare, A. J. (2001). The neuroendocrinology of chronic fatigue syndrome and fibromyalgia. *Psychological Medicine*, 31, 1331-1345.
- Parving, A. & Christensen, B. (1993). Training and employment in hearing-impaired subjects at 20-35 years of age. *Scandinavian Audiology*, 22, 133-139.
- Paulus, M. P., Rogalsky, C., Simmons, A., Feinstein, J. S. & Stein, M. B. (2003). Increased activation in the right insula during risk-taking decision making is related to harm avoidance and neuroticism. *Neuroimage*, 19, 1439-1448.

- Paykel, E. S. (2001). Stress and affective disorders in humans, *Seminars in Clinical Neuropsychiatry*, 6, 4-11.
- Peelle, J. E. (2017). Listening effort: How the cognitive consequences of acoustic challenge are reflected in brain and behaviour. *Ear & Hearing*, 39, 204-214.
- Peelle, J. E., Troiani, V., Wingfield, A. & Grossman, M. (2010). Neural processing during older adults' comprehension of spoken sentences: Age differences in resource allocation and connectivity. *Cerebral Cortex*, 20, 773-782.
- Perreau, A. E., Wu, Y. -H., Tatge, B., Irwin, D. & Corts, D. (2017). Listening effort measured in adults with normal hearing and cochlear implants. *Journal of the American Academy of Audiology*, 28, 685-697.
- Peterson, N. R., Pisoni, D .B. & Miyamoto, R. T. (2010). Cochlear implants and spoken language processing abilities: review and assessment of the literature. *Restorative Neurology and Neuroscience*, 28, 237-250.
- Phillips, N. (2016). The implications of cognitive aging for listening and the FUEL model. *Ear & Hearing*, 37, 44S-51S.
- Phonak (2010). *Dynamic SoundField: The technology behind the world's leading soundfield system*. pp.1-12.
- Pichora-Fuller, M. K. (2016). How social psychological factors may modulate auditory and cognitive functioning during listening. *Ear & Hearing*, 37, 92S-100S.
- Pichora-Fuller, M. K. & Kramer, S. E. (2016). Eriksholm Workshop on Hearing Impairment and Cognitive Energy. *Ear & Hearing*, 37, 1S-4S.
- Pichora-Fuller, M. K., Johnson, C. E. & Roodenburg, K. E. J. (1998). The discrepancy between hearing impairment and handicap in the elderly: Balancing transaction and interaction in conversation. *Journal of Applied Communication Research*, 26, 99-119.
- Pichora-Fuller, M. K., Mick, P. T. & Reed, M. (2015). Hearing, cognition and healthy aging: Social and public health implications of the links between age-related declines in hearing and cognition. *Seminars in Hearing*, 36, 122-139.
- Pichora-Fuller, M. K., Kramer, S. E., Eckert, M. A., Edwards, B., Hornsby, B. W. Y., Humes, L. E., Lemke, U., Lunner, T., Matthen, M., Mackersie, C. L., Naylor, G., Phillips, N. A., Richter, M., Rudner, M., Sommers, M. S., Tremblay, K. L. & Wingfield, A. (2016). Hearing impairment and cognitive energy: The framework for understanding effortful listening (2016), *Ear & Hearing*, 37, 5S-27S.
- Picou, E. (2013). 20Q: Listening Effort – We know it's a problem, but how do we measure it? *AudiologyOnLine*. Retrieved from <http://www.audiologyonline.com/articles/listening-effort-we-know-it-12032>

- Picou, E. M., Ricketts, T. A. & Hornsby, B. W. (2011). Visual cues and listening effort: individual variability. *Journal of Speech, Language and Hearing Research*, 54, 1416-1430.
- Picou, E. M., Ricketts, T. A. & Hornsby, B. W. (2013). How hearing aids, background noise and visual cues influence objective listening effort, *Ear & Hearing*, 34, e52-e64.
- Pine, D. S., Cohen, P. & Brook, J. S. (2001). Emotional reactivity and risk for psychopathology among adolescents. *CNS Spectrums*, 6, 27-35.
- Piquado, T., Issacowitaz, D. & Wingfield, A. (2010). Pupillometry as a measure of cognitive effort in younger and older adults. *Psychophysiology*, 47, 560-569.
- Pisoni, D. B., Cleary, M., Geers, A. E. & Tobey, E. A. (1999). Individual differences in effectiveness of cochlear implants in children who are prelingually deaf: New process measures of performance. *Volta Review*, 101, 111-164.
- Plomp, R., & Mimpen, A. (1979). Improving the reliability of testing the speech reception threshold for sentences. *Audiology*, 18, 43–52.
- Poissant, S. F., Bero, E. M., Busekroos, L. & Weiru, S. (2014). Determining cochlear implants users' true noise tolerance: Use of speech reception threshold in noise testing. *Otology & Neurotology*, 34, 414-420.
- Polich, J. (2003). Overview of the P3a and P3b. In J. Polich (Ed.), *Detection of change: Event-related potential and fMRI findings* (pp. 83-98). Boston, MA: Kluwer Academic Press.
- Postle, B. R. (2006). Working memory as an emergent property of the mind and brain. *Neuroscience*, 139, 23-38.
- Potter, M. (1976). Short-term conceptual memory for pictures. *Journal of Experimental Psychology: Human Learning and Memory*, 2, 509-522.
- Potter, M. (1993). Very short-term conceptual memory. *Memory and Cognition*, 21, 156-161.
- Provenzano, M. J. & Domann, F. E. (2007). A role for epigenetics in hearing: Establishment and maintenance of auditory specific gene expression patterns. *Hearing Research*, 233, 1-13.
- Qin, M. K., & Oxenham, A. J. (2003). Effects of simulated cochlear-implant processing on speech reception in fluctuating maskers. *Journal of the Acoustical Society of America*, 114, 446–454.
- Rabbitt, P. M. (1968). Channel-capacity, intelligibility and immediate memory. *Quarterly Journal of Experimental Psychology*, 20, 241-248.
- Rabbitt, P. M. (1990). Mild hearing loss can cause apparent memory failures which increase with age and reduce with IQ. *Acta Oto-Laryngologica Supplement*, 476, 167-176.

References

- Radley, J. J., Rocher, A. B., Rodriguez, A., Ehlenberger, D. B., Dammann, M., McEwen, B. S., Morrison, J. H., Wearne, S. I. & Hof, P. R. (2008). Repeated stress alters dendritic morphology in the rat medial prefrontal cortex. *Journal of Comparative Neurology*, 507, 1141-1150.
- Rakerd, B., Seitz, P. F. & Whearty, M. (1996). Assessing the cognitive demands of speech listening for people with hearing losses. *Ear & Hearing*, 17, 97-106.
- Rees, G., Russell, C., Frith, C. D., Driver, J. (1999). Inattention blindness versus inattentional amnesia for fixated but ignored words. *Science*, 286, 2504-2507.
- Rennies, J., Schepker, H., Holube, I. & Kolmeier, B. (2014). Listening effort and speech intelligibility in listening situations affected by noise and reverberation. *Journal of the Acoustical Society of America*, 136, 2642-2653.
- Ribeiro, S., Goyal, V., Mello, C. V. & Pavlides, C. (1999). Brain gene expression during REM sleep depends on prior waking experience. *Learning & Memory*, 6, 500-508.
- Richter, M. (2016). The moderating effect of success importance on the relationship between listening demand and listening effort, *Ear & Hearing*, 37, 111S-117S.
- Richter, M., Gendolla, G. H. E. & Wright, R. A. (2016). Three decades of research on motivational intensity theory: What we have learned about effort and what we still don't know. *Advances in Motivation Sciences*, 3, 149-186.
- Roberts, A. D. L., Charler, M. -L., Papadopoulos, A., Wessely, S., Chalder, T. & Cleare, A. J. (2009). Does hypocortisolism predict a poor response to cognitive behavioural therapy in chronic fatigue syndrome? *Psychological Medicine*, 40, 515-522.
- Robles, L. & Ruggero, M. A. (2001). Mechanics of the mammalian cochlea. *Physiological Reviews*, 81, 1305-1352.
- Rönnberg, J. (2003). Cognition in the hearing impaired and deaf as a bridge between signal and dialogue: A framework and a model. *International Journal of Audiology*, 42, 68-76.
- Rönnberg, J., Arlinger, S., Lyxell, B., & Kinnefors, C. (1989). Visual evoked potentials: Relations to adult speechreading and cognitive function. *Journal of Speech and Hearing Research*, 32, 725-735.
- Rönnberg, J., Rudner, M., Food, C. & Lunner, T. (2008). Cognition counts: A working memory system for ease of language understanding (ELU). *International Journal of Audiology*, 47, S99-S105.
- Rönnberg, J., Lunner, T., Zekveld, A., Sörqvist, P., Danielsson, H., Lyxell, B., Dahlström, Signoret, C., Stenfelt, S., Pichora-Fuler, M. K. & Rudner, M. (2013). The ease of language understanding (ELU) model: Theoretical, empirical, and clinical advances. *Frontiers in Systems Neuroscience*, 7, 1-17.

- Rönnberg, J., Rudner, M. & Lunner, T. (2014). Comments: Listening effort and fatigue. *International Journal of Audiology*, 53, 433-445.
- Rosen, S. (2011). *The complexities of understanding speech in background noise*, First International Conference on Cognitive Hearing Science for Communication, 19 June, Linköping, Sweden.
- Rosen, S., Faulkner, A. & Wilkinson, L. (1999). Adaptation by normal listeners to upward spectral shifts of speech: Implications for cochlear implants. *Journal of the Acoustical Society of America*, 106, 3629-3636.
- Roy, A., Pickar, D., De Jong, J., Karoum, E. & Linnoila, M. (1988). Norepinephrine and its metabolites in cerebrospinal fluid, plasma and urine. Relationship to hypothalamic-pituitary-adrenal axis function in depression. *Archives of General Psychiatry*, 45, 849-857.
- Rubenstein, J. T. (2012). Cochlear implants: the hazards of unexpected success. *Canadian Medical Association Journal*, 184, 1343-1344.
- Rudner, M. (2016). Cognitive spare capacity as an index of listening effort. *Ear & Hearing*, 37, 69S-76S.
- Rudner, M., Foo, C., Rönnberg, J. & Lunner, T. (2009). Cognition and aided speech recognition in noise: Specific role for cognitive factors following nine-week experience with adjusted compression settings in hearing aids. *Scandinavian Journal of Psychology*, 50, 405-418.
- Rudner, M., Davidsson, L. & Rönnberg, J. (2010). Effects of age on the temporal organization of working memory in deaf signers. *Neuropsychology and Developmental Cognition B in Aging Neuropsychology and Cognition*, 17, 360-383.
- Rudner, M., Lunner, T., Behrens, T., Thoréns, E.S. & Rönnberg, J. (2012). Working memory capacity may influence perceived effort during aided speech recognition in noise. *Journal of the American Academy of Audiology*, 23, 577-589.
- Ryan, E. B. & Giles, H., Bartolucci, G. & Henwood, K. (1986). Psycholinguistic and social psychological components of communication by and with the elderly. *Language and Communication*, 6, 1-24.
- Ryan, E. B., Meredith, S. D., MacLean, M. J. & Orange, J. B. (1995). Changing the way we talk with elders: Promoting health using the communication enhancement model. *International Journal of Aging and Human Development*, 41, 89-107.
- Sandi, C. & Haller, J. (2015). Stress and the social brain: Behavioural effects and neurobiological mechanisms. *Nature Reviews Neuroscience*, 16, 290-304.

- Sarampalis, A., Sridhar, K., Edwards, B. & Hafter, E. (2009). Objective measures of listening effort: Effects of background noise and noise reduction. *Journal of Speech, Language and Hearing Research*, 52, 1230-1240.
- Sauseng, P. & Klimesch, W. (2008). What does phase information of oscillatory brain activity tell us about cognitive processes? *Neuroscience and Biobehavioural Reviews*, 32, 1001-1013.
- Schafer, E. C. & Thibodeau, L. M. (2004). Speech recognition abilities of adults using cochlear implants with FM systems. *Journal of the American Academy of Audiology*, 15, 678-791.
- Schlotz, W., Hellhammer, J., Schulz, P. & Stone, A. A. (2004). Perceived work overload and chronic worrying predict weekend-weekday differences in the cortisol awakening response. *Psychosomatic Medicine*, 66, 207-214.
- Schneiderman, N. (1983). Pathophysiology in animals. In T. M. Dembroski, T. H. Schmidt & G. Blümhen (Eds.), *Biobehavioral Bases of Coronary Heart Disease* (pp. 304-364). Basel, Germany: Karger.
- Schneiderman, N., Ironson, G. & Siegel, S. D. (2005). Stress and health: Psychological, behavioural, and biological determinants. *Annual Review of Clinical Psychology*, 1, 607-628.
- Schoof, T. & Rosen, S. (2014). The role of auditory and cognitive factors in understanding speech in noise by normal-hearing older listeners. *Frontiers in Aging Neuroscience*, 6, 1-14.
- Schork, M. & Williams, G. (1980). Number of observations required for the comparison of two correlated proportions. *Communications in Statistics-Simula.Computer*, B9, 349-356.
- Schramm, D., Fitzpatrick, E. & Seguin, C. (2002). Cochlear implantation for adolescents and adults with prelinguistic deafness. *Otology & Neurotology*, 23, 698-703.
- Segerstrom, S. C. & Miller, G. E. (2004). Psychological stress and the human immune system: a meta-analytic study of 30 years of inquiry. *Psychological Bulletin*, 130, 601-630.
- Selye, H. (1956). *The stress of life*. New York, NY: McGraw-Hill.
- Shannon, R. V., Zeng, F. –G., Kamath, V., Wygonski, J. & Ekelid, M. (1995). Speech recognition with primarily temporal cues. *Science*, 270, 303-304.
- Shannon, R. V., Zeng, F. –G. & Wygonski, J. (1998). Speech recognition with altered spectral distribution of envelope cues. *Journal of the Acoustical Society of America*, 104, 2467-2476.

- Shannon, R. V., Jensvold, A., Padilla, M., Robert, M. E. & Wang, X. (1999). Consonant recordings for speech testing. *Journal of the Acoustical Society of America*, 106, L71-L74.
- Sharma, A., Gilley, P. M., Dorman, M. F. & Baldwin, R. (2007). Deprivation-induced cortical reorganization in children with cochlear implants. *International Journal of Audiology*, 46, 494-499.
- Sharma, A., Nash, A. A. & Dorman, M. (2009). Cortical development, plasticity and reorganization in children with cochlear implants. *International Journal of Language and Communication Disorders*, 42, 272-279.
- Shepherd, R. K. & Javel, E. (1997). Electrical stimulation of the auditory nerve. I. Correlation of physiological responses with cochlear status. *Hearing Research*, 108, 112-144.
- Sherrington, C. S. (1897). On the reciprocal action in the retina as studied by means of some rotating discs. *Journal of Physiology*, 21, 33-54.
- Shimojo, S., Kamitani, Y. & Nishida, S. (2001). Afterimage of perceptually filled-in surface. *Science*, 293, 1677-1680.
- Shinn-Cunningham, B. G. & Best, V. (2008). Selective attention in normal and impaired hearing. *Trends in Amplification*, 12, 283-299.
- Silveri, M. M., Tzilos, G. K., Pimentel, P. J. & Yurgelun-Todd, D. A. (2004). Trajectories of adolescent emotional and cognitive development: effects of sex and risk for drug use. *Annals of the New York Academy of Sciences*, 1021, 363-370.
- Sirios, S. & Brisson, J. (2014) Pupillometry. *Wiley Interdisciplinary Reviews: Cognitive Science*, 5, 679-692.
- Sisk, C. I. & Foster, D. L. (2004). The neural basis of puberty and adolescence. *Nature Neuroscience*, 7, 1040-1047.
- Skinner, M. W., Ketten, D. R., Holden, L. K., Harding, G. W., Smith, P. G., Gates, G. A., Neely, J. G., Kletzker, G. R., Brunsden, B. & Blocker, B. (2002). CT-derived estimation of cochlear morphology and electrode array position in relation to word recognition in Nucleus-22 recipients. *Journal of the Association for Research in Otolaryngology*, 3, 332-350.
- Skinner, M. W., Holden, T. A., Whiting, B. R., Voie, A. H., Brunsden, B., Neely, J. G., Saxon, E. A., Hullar, T. E. & Finley, C. C. (2007). In vivo estimates of the position of advanced bionics electrode arrays in the human cochlea. *Annals of Otolaryngology, Rhinology & Laryngology*, 116, 2-24.
- Smith, R. L. (1985). Cochlear processes reflected in responses of the cochlear nerve. *Acta Otolaryngologica*, 100, 1-12.

- Smith, Z. M. & Delgutte, B. (2008). Sensitivity of inferior colliculus neurons to interaural time differences in the envelope versus the fine structure with bilateral cochlear implants. *Journal of Neurophysiology*, 99, 2390-2407.
- Smith, Z. M., Delgutte, B. & Oxenham, A. J. (2002). Chimaeric sounds reveal dichotomies in auditory perception. *Nature*, 416, 87-90.
- Southall, K., Gagné, J. -P. & Jennings, M. B. (2010). Stigma: A negative and a positive influence on help-seeking for adults with acquired hearing loss. *International Journal of Audiology*, 49, 804-814.
- Spear, L.P. (2000). The adolescent brain and age-related behavioural manifestations. *Neuroscience & Biobehavioural Reviews*, 24, 417-463.
- Spence, R. (2002). Rapid, serial and visual: A presentation technique with potential. *Information Visualization*, 1, 13-19.
- Spence, R. & Witkowski, M. (2013). *Rapid serial visual presentation*. Design for Cognition, SpringerBriefs in Computer Science: Springer.
- Sporns, O. (2015). Cerebral cartography and connectomics. *Philosophical Transactions of the Royal Society London Biological Sciences*, 370, 1-12.
- Spruill, T. M. (2010). Chronic psychosocial stress and hypertension. *Current Hypertension Reports*, 12, 10-16.
- Srinivasan, A.G., Padilla, M., Shannon, R. V. & Landsberger, D. M. (2013). Improving speech perception in noise with current focusing in cochlear implant users. *Hearing Research*, 299, 29-36.
- Steacie, P., Olds, J., Shaw-Pickard, E., Whittingham, J. -A., Neuss, D., McSweeney, E., Rabjohn, K., Somerville, R., Beauregard, Y. & Schramm, D. (2016). *Integrating CI sound processor data logging into auditory-verbal practice: An examination of its usefulness for practitioners and families*. Retrieved from https://cdn.ymaws.com/www.acialliance.org/resource/resmgr/CI2016_Abstracts/ci_data_logging_powerpoint_w.pdf.
- Steel, M. M., Papsin, B. C. & Gordon, K. A. (2015). Binaural fusion and listening effort in children who use bilateral cochlear implants: A psychoacoustic and pupillometric study. *PLOS ONE*, 1-29.
- Steinberg, L. (2005). Cognitive and affective development in adolescence. *Trends in Cognitive Sciences*, 9, 69-74.
- Steiner, J., Mawrin, C., Ziegeler, A., Bielau, H., Ullrich, O., Bernstein, H. -G. & Bogerts, B. (2006). Distribution of HLA-DR-positive microglia in schizophrenia reflects impaired cerebral lateralization. *Acta Neuropathologica*, 112, 305-316.

- Steiner, J., Bielau, H., Brisch, R., Danos, P., Ullrich, O., Mawrin, C., Bernstein, H. –G. & Bogerts, B. (2008). Immunological aspects in the neurobiology of suicide: elevated microglial density in schizophrenia and depression is associated with suicide. *Journal of Psychiatric Research*, 42, 151-157.
- Stephens, M. A. C. & Wand, G. (2012). Stress and the HPA Axis: Role of glucocorticoids in alcohol dependence. *Alcohol Research: Current Reviews*, 468-483.
- Stickney, G. S., Loizou, P. C., Mishra, L. N., Assmann, P. F., Shannon, R. V. & Opie, J. M. (2006). Effects of electrode design and configuration on channel interactions. *Hearing Research*, 211, 33-45.
- Stineman, M. G. & Streim, J. E. (2010). The biopsychosocial-ecological paradigm: A foundational theory for medicine. *Physical Medicine & Rehabilitation*, 2, 1035-1045.
- Stokols, D. (1992). Establishing and maintaining healthy environments. Towards a social ecology of health promotion. *American Psychologist*, 47, 6-22.
- Stone, M. A. & Moore, B. C. J. (2003). Effect of the speed of a single-channel dynamic range compressor on intelligibility in a competing speech task. *The Journal of the Acoustical Society of America*, 114, 1023-1034.
- Strawbridge, W. J., Wallhagen, M. I., Shema, S. J. & Kaplan, G. A. (2000). Negative consequences of hearing impairment in old age: A longitudinal analysis. *The Gerontologist*, 40, 320-326.
- Styles, E. A. (2006). *The Psychology of Attention* (2nd ed.). New York, NY: Oxford University Press.
- Suprenant, A. M. (2007). Effects of noise on identification and serial recall of nonsense syllables in older and younger adults. *Developmental Cognition B in Aging Neuropsychology and Cognition*, 14, 126-143.
- Surowiecki, V. N., Maruff, P., Busby, P. A., Sarant, J., Blarney, P. J. & Clark, G. M. (2002). Cognitive processing in children using cochlear implants: The relationship between visual memory, attention, and executive functions and developing language skills. *Annals of Otology, Rhinology & Laryngology*, 111, 119–126.
- Svirsky, M. A., Silveira, A., Neuburger, H., Teoh, S. W. & Suarez, H. (2004). Long-term auditory adaptation to a modified peripheral frequency map. *Acta Oto-Laryngologica*, 124, 381-386.
- Technavio. (2016). *The number of cochlear implants shipped worldwide is expected to hit 96,000 by 2020*. Retrieved from <https://www.technavio.com/blog/number-cochlear-implants-shipped-worldwide-expected-hit-96000-2020>

- Thomas, K. M., Hunt, R. H., Vizueta, N., Sommer, T., Durston, S., Yang, Y. & Worden, M. S. (2004). Evidence of developmental differences in implicit sequence learning: an fMRI study of children and adults. *Journal of Cognitive Neuroscience*, 16, 1339-1351.
- Tillein, J., Hartmann, R. & Kral, A. (2015). Electric-acoustic interactions in the hearing cochlea: Single fiber recordings. *Hearing Research*, 322, 112-126.
- Titchener, E. B. (1908). *Lectures on the Elementary Psychology of Feeling and Attention*. New York, NY: MacMillan.
- Tobey, E. A., Thal, D., Niparko, J. K., Eisenberg, L. S., Quittner, A. L., Wang, N. -Y. & Team, C. D. I. (2013). Influence of implantation age on school-age language performance in pediatric cochlear implant users. *International Journal of Audiology*, 52, 219-229.
- Tong, L., Strong, M. K., Kaur, T., Juiz, J. M., Oesterle, E. C., Hume, C., Warchol, M. E., Palmiter, R. D. & Rubel, E. W. (2015). Selective depletion of cochlear hair cells cause rapid age-dependent changes in spiral ganglion and cochlear nucleus neurons. *Journal of Neuroscience*, 35, 7878-7891.
- Tremblay, K. & Backer, K. (2016). Listening and learning: Implications for the rehabilitation of adults with and without hearing loss. *Ear & Hearing*, 37, 155S-162S.
- Uhlmann, R. F., Larson, E. B., Rees, T. S., Koepsell, T. D. & Duckert, L. G. (1989). Relationship of hearing impairment to dementia and cognitive dysfunction in older adults. *Journal of the American Medical Association*, 261, 1916-1919.
- Ulrich, R. & Miller, J. (2004). Threshold estimation in two-alternative forced-choice (2AFC) tasks: The Spearman-Kärber method. *Perception & Psychophysics*, 66, 517-533.
- Ulrich-Lai, Y. M. & Herman, J. P. (2009). Neural regulation of endocrine and autonomic stress responses. *Nature Reviews Neuroscience*, 10, 397-409.
- Vaden, K. I. Jr., Kuchinsky, S. E., Cute, S. L., Ahlstrom, J. B., Dubno, J. R. & Eckert, M. A. (2013). The cingulo-opercular network provides word-recognition benefit. *Journal of Neuroscience*, 33, 18979-18986.
- Vaden, K. I. Jr., Kuchinsky, S. E., Ahlstrom, J. B., Dubno, J. R. & Eckert, M. A. (2015). Cortical activity predicts which older adults recognize speech in noise and when. *Journal of Neuroscience*, 35, 3929-3937.

- van Berckel, B. N., Bossong, M. G., Boellaard, R., Kloet, R., Schuitmaker, A., Caspers, E., Luurtsema, G., Windhorst, A. D., Cahn, W., Lammertsma, A. & Kahn, R. S. (2008). Microglia activation in recent-onset schizophrenia: a quantitative (R)-[11C]PK11195 positron emission tomography study. *Biological Psychiatry*, 64, 820-822.
- Van den Honert, C. & Stypulkowski, P. H. (1987). Single fiber mapping of spatial excitation patterns in the electrically stimulated auditory nerve. *Hearing Research*, 20, 195-206.
- van Esch, T. E. M., Kollmeier, B., Vormann, M., Lyzenga, J., Houtgast, T., Hällgren, M., Larsby, B., Athalye, S. P., Lutman, M. E. & Dreschler, W. A. (2013). Evaluation of the preliminary auditory profile test battery in an international multi-centre study. *International Journal of Audiology*, 52, 305-321.
- Van Orden, K. F., Jung, T. -P. & Makeig, S. (2000). Combined eye activity measures accurately estimate changes in sustained visual performance. *Biological Psychology*, 52, 221-240.
- Varazzani, C., San-Galli, A., Gilardeau, S. & Bouret, S. (2015). Noradrenaline and dopamine neurons in the reward/effort trade-off: A direct electrophysiological comparison in behaving monkeys. *Journal of Neuroscience*, 35, 7866-7877.
- Vertes, R. P. (2004). Memory consolidation in sleep: Dream or reality. *Neuron*, 44, 135-148.
- Vidlock, E. J., Shih, W., Adeyemo, M., Mahurkar-Joshi, S., Presson, A. P., Polytarchou, C., Alberto, M., Iliopoulos, D., Mayer, E. A. & Chang, L. (2016). The effect of sex and irritable bowel syndrome on HPA axis response and peripheral glucocorticoid receptor expression. *Psychoneuroendocrinology*, 69, 67-76.
- Vorberg, D., Mattler, U., Heinecke, A., Schmidt, T. & Schwarzbach, J. (2003). Different time courses for visual perception and action priming. *Proceedings of the National Academy of Sciences of the United States of America*, 100, 6275-6280.
- Walker, M. P. (2008). Cognitive consequences of sleep and sleep loss. *Sleep Medicine*, 9, 29S-34S.
- Wallis, C. (2006, March 27). The multitasking generation. *Time*, 163.
- Walsh, B. J., Buonocore, M. H., Carter, C. S. & Mangun, G. R. (2011). Integrating conflict detection and attentional control mechanisms. *Journal of Cognitive Neuroscience*, 23, 2211-2221.

- Wass, M., Ibertsson, T., Lyxell, B., Sahlen, B., Hällgren, M., Larsby, B. & Mäki-Torkko, E. (2008). Cognitive and linguistic skills in Swedish children with cochlear implants: measures of accuracy and latency as indicators of development. *Scandinavian Journal of Psychology*, 49, 559-576.
- Watson, J. M. & Strayer, D. L. (2010). Supertaskers: Profiles in extraordinary multitasking ability. *Psychonomic Bulletin & Review*, 17, 479-485.
- Wechsler, D (1999). *Wechsler Abbreviated Scale of Intelligence*. San Antonio, Texas: Pearson Corporation.
- Weichselgartner, E. & Sperling, G. (1987). Dynamics of automatic and controlled visual attention. *Science*, 238, 778-780.
- Weinberger, N. M. (2011). The medial geniculate, not the amygdala, as the root of auditory fear conditioning. *Hearing Research*, 274, 61-74.
- Weinstein, B. E. & Ventry, I. M. (1982). Hearing impairment and social isolation in the elderly. *Journal of Speech and Hearing Research*, 24, 593-599.
- Weisz, N., Hartmann, T., Müller, N., Lorenz, I. & Obleser, J. (2011). Alpha rhythms in audition: Cognitive and clinical perspectives. *Frontiers in Psychology*, 2, 73.
- Whitehead, D. L., Perkins-Porras, L., Strike, P. C., Magid, K. & Steptoe, A. (2007). Cortisol awakening response is elevated in acute coronary syndrome patients with type-D personality. *Journal of Psychosomatic Research*, 62, 419-425.
- Whiteus, C., Freitas, C. & Grutzendler, J. (2014). Perturbed neural activity disrupts cerebral angiogenesis during a postnatal critical period. *Nature*, 505, 407-411.
- WHO (2017). *Factsheet: Deafness and hearing loss*. Retrieved from <http://www.who.int/mediacentre/factsheets/fs300/en/>
- Wichmann, F. A. & Hill, N. J. (2001). The psychometric function: I. Fitting, sampling and goodness of fit. *Perception & Psychophysics*, 63, 1293-1313.
- Wiggins, I.M., Anderson, C. A., Kitterick, P. T. & Hartley, D. E. H. (2016). Speech-evoked activation in adult temporal cortex measured using functional near-infrared spectroscopy (fNIRS): Are the measurements reliable? *Hearing Research*, 339, 142-154.
- Wijayasari, P., Hartley, D. E. H. & Wiggins, I. M. (2017). Brain activity underlying the recovery of meaning from degraded speech: A functional near-infrared spectroscopy (fNIRS) study. *Hearing Research*, 351, 55-67.
- Wild, C. J., Yusuf, A., Wilson, D. E., Peelle, J. E., Davis, M. H. & Johnsrude, I. S. (2012). Effortful listening: The processing of degraded speech depends critically on attention. *The Journal of Neuroscience*, 32, 14010-14021.
- Wilson, B. S. (1997). The future of cochlear implants. *British Journal of Audiology*, 31, 205-225.

- Wilson, B. S. (2015). Getting a decent (but sparse) signal to the brain for users of cochlear implants. *Hearing Research*, 322, 24-38.
- Wilson, B. S. & Dorman, M. F. (2008a). Cochlear implants: Current designs and future possibilities. *Journal of Rehabilitation Research & Development*, 45, 695-730.
- Wilson, B. S. & Dorman, M. F. (2008b). Cochlear implants: A remarkable past and a brilliant future. *Hearing Research*, 242, 3-21.
- Wilson, B. S., Schatzer, R., Lopez-Poveda, E. A., Sun, X., Lawson, D. T. & Wolford, R. D. (2005). Two new directions in speech processor design for cochlear implants. *Ear & Hearing*, 26, 73S-81S.
- Wilson, B. S., Dorman, M. F., Woldorff, M. G. & Tucci, D. L. (2011). Cochlear implants matching the prosthesis to the brain and facilitating desired plastic changes in brain function. *Progress in Brain Research*, 194, 117-129.
- Wingfield, A., McCoy, S. L., Peelle, J. E., Tun, P. A. & Cox, L. C. (2006). Effects of adult aging and hearing loss on comprehension of rapid speech varying in syntactic complexity. *Journal of the American Academy of Audiology*, 17, 487-497.
- Wingfield, A. (2016). Evolution of working memory and cognitive resources. *Ear & Hearing*, 37, 35S-43S.
- Winn, M. B., Edwards, J. R. & Litovsky, R. Y. (2015). The impact of auditory spectral resolution on listening effort revealed by pupil dilation. *Ear & Hearing*, 36, 153-165.
- Worrall, L. E. & Hickson, L. M. (2003). *Communication disability in aging: From prevention to intervention*. New York, NY: Delmar Learning.
- Wright, R. A., Gollwitzer, P. M. & Bargh, J. A. (1996). Brehm's theory of motivation as a model of effort and cardiovascular response. In *The Psychology of Action: Linking Cognition and Motivation to Behaviour* (pp.424-453). New York, NY: Guilford.
- Wu, Y. -H., Stangl, E., Zhang, X., Perkins, J. & Eilers, E. (2016). Psychometric functions of dual-task paradigms for measuring listening effort. *Ear & Hearing*, 37, 660-667.
- Yang, Z. & Cossetti, M. (2016). Safety and outcomes of cochlear implantation in the elderly: A review of the recent literature. *Journal of Otology*, 11, 1-6.
- Ye, H., Deng, G., Mauger, S. J., Hersbach, A. A., Dawson, P. W. & Heasman, J. M. (2013). A wavelet-based noise reduction algorithm and its clinical evaluation in cochlear implants. *PLOS ONE*, 9, e75662.
- Yerkes, R. M. & Dodson, J. D. (1908). The relation of strength of stimulus to rapidity of habit-formation. *Journal of Comparative Neurology and Psychology*, 18, 459-482.

- Yukawa, K., Cohen, L., Blamey, P., Pyman, B., Tungvachirakul, V. & O'Leary, S. (2003). Effects of insertion depth of cochlear implant electrodes upon speech perception. *Audiology & Neurotology*, 9, 163-172.
- Zec, R. F. (1995). Neuropsychology of schizophrenia according to Kraepelin: disorders of volition and executive functioning. *European Archives of Psychiatry and Clinical Neuroscience*, 245, 216-223.
- Zekveld, A. A., & Kramer, S. E. (2014). Cognitive processing load across a wide range of listening conditions: Insights from pupillometry. *Psychophysiology*, 51, 277–284.
- Zekveld, A. A, Kramer, S. E., & Festen, J. M. (2010). Pupil response as an indication of effortful listening: the influence of sentence intelligibility. *Ear & Hearing*, 31, 480–490.
- Zekveld, A. A, Kramer, S. E., & Festen, J. M. (2011). Cognitive load during speech perception in noise: The influence of age, hearing loss, and cognition on the pupil response. *Ear & Hearing*, 32, 498–510.
- Zekveld, A. A., Heslenfeld, D. J., Johnsrude, I. S., Versfeld, N. J., & Kramer, S. E. (2014a). The eye as a window to the listening brain: Neural correlates of pupil size as a measure of cognitive listening load. *NeuroImage*, 101, 76–86.
- Zekveld, A. A., Rudner, M., Kramer, S. E., Lyzenga, J., & Rönnberg, J. (2014b). Cognitive processing load during listening is reduced more by decreasing voice similarity than by increasing spatial separation between target and masker speech. *Frontiers in Neuroscience*, 8, 1–11.
- Zeng, F. G. (2002). Temporal pitch in electric hearing. *Hearing Research*, 174, 101-106.
- Ziegler, D. R., Cullinan, W. E. & Herman, J. P. (2005). Organization and regulation of paraventricular nucleus glutamate signalling systems: N-methyl-D-aspartate receptors. *Journal of Comparative Neurology*, 484, 43-56.